The Thesis committee for Peter Lisle Schaeffer
Certifies that this is the approved version of the following thesis:

A Simple Sub-1V Voltage Reference

APPROVED BY
SUPERVISING COMMITTEE:

T.R. Viswanathan, Supervisor

Deji Akinwande

A Simple Sub-1V Voltage Reference

by

Peter Lisle Schaeffer

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

December 2017

Dedication

Dedicated to my wife and my Mom and Dad.

Acknowledgements

First, I would like to thank my loving wife for supporting me so much while I finish my Masters degree and this thesis. She has sacrificed so much for me and taken care of me so well, and I am forever thankful for the love she has shown me. I thank my Mom and my Dad for working so hard to raise me and my four siblings, and for all the sacrifices they have made for us. I thank them for preparing me for university and life and for all of their love. I thank my Lord and Savior, Jesus Christ, for His love and sacrifice on the cross so that I may have life in Him.

I would like to thank my supervisor, T.R. Viswanathan, for encouraging me to work on this thesis and for all of his help along the way. I greatly enjoyed all of our meetings and hearing so many great stories from his years of experience. I thank K.R. Raghunandan for his work on the tape-out for the voltage reference circuit, and for all of his support throughout my simulation and silicon testing phase. I thank Deji Akinwande for serving as a reader for this thesis.

I am very grateful for Professor Nan Sun in teaching both the introductory analog IC design course and also the data converters class. Through him, I gained exposure to basic and advanced topics in analog IC design and his courses helped to develop my strong passion in the field. Furthermore, I thank him for allowing me to take the measurements in his lab in UT Austin.

I thank all of my friends I have made during my time at UT Austin. I'm thankful for my friends I made from being in Bridges International at UT, with a special thank you to Kathrine and Adriana Yauw and Sarah Moreno, for being so welcoming to me when I was a freshman and serving in Bridges for three years together. We had so many great

experiences and my college experience (and my life) would be much different without them. I also want to thank my senior design team, who are also great friends: Ryan Chow, Tu-An Nguyen, Wallace Tran, Derek Young, and Wen Zhang. All of them made senior design a great experience and allowed me take my RFIC course during one of the lab meeting times. They all made my time in university a truly great and memorable experience.

Abstract

A Simple Sub-1V Voltage Reference

Peter Lisle Schaeffer, MSE

The University of Texas at Austin, 2017

Supervisor: T.R. Viswanathan

A traditional bandgap reference voltage has a value approximately equal to 1.2V. This is an inconvenient value to obtain in short-channel CMOS circuits where the nominal power supply voltage is 1.2V or lower. Here, a simple circuit is presented to generate a reference voltage between 600mV and 800mV, based on the threshold voltage of a MOS device, which will be suitable for biasing transistors in strong inversion as well as for use in 6-bit data converters needed for high speed data communication systems. The circuit uses a novel, but simple, method for generating the temperature coefficients needed for generating a stable voltage reference across temperature. The core circuit produces a PTAT current that when passed through a diode-connected device, with a CTAT threshold voltage, will produce a voltage that is stable across temperature. The measured results demonstrate an average voltage variation of 4.62mV across five chips, each containing eight voltage reference circuits, and a minimum voltage variation of 1.1mV from 0°C to 80°C. A variation of the proposed circuit replaces the BJTs with MOSFETs to achieve smaller area, and simulation results indicate even less voltage variation across temperature.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 MOTIVATION

Modern-day integrated circuits (ICs) are used in a variety of environments, which can expose ICs to a wide-range of temperatures. Additionally, the die of the IC itself can heat up due to power dissipation in the circuit. Electronic components' parameters change as temperature changes. Resistors, V_{BE}, in bipolar transistors, and V_{th}, the threshold voltage in MOSFETs, are all examples of components that vary with temperature. Thus, the circuit performance will vary with temperature. As an application-specific example, for a 6-bit ADC using a 1V full-scale input, the LSB is approximately 16mV. Thus, if the temperature variation causes a voltage variation large enough to affect the levels in an ADC, then the comparison reference voltage variation will cause an error in the quantization of the ADC.

The bandgap reference circuit, first introduced by Bob Widlar in 1971, is a fundamental circuit now commonly used in most analog integrated circuits (ICs). The circuit's main purpose is to provide a reference voltage and current that will be constant across all temperatures. This helps absorb variation introduced by temperature variation. The fundamental operation of any stable voltage reference across temperature is to combine two temperature-dependent terms, one being proportional to absolute temperature (PTAT), and the other being complementary to absolute temperature (CTAT). This is illustrated in Figure 1.1. By adding these two terms together, a reference voltage, V_{REF}, is generated, which will be flat across temperature.

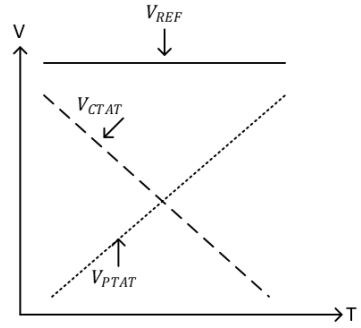

Figure 1.1: Ideal Voltage Reference Generation

In practice, there are higher order terms which will make a CTAT quantity decrease non-linearly with temperature. This is shown in figure 1.2. The reference voltage at the output will therefore not be constant across temperature because of the non-linearity. This is referred to as curvature and needs to be compensated.

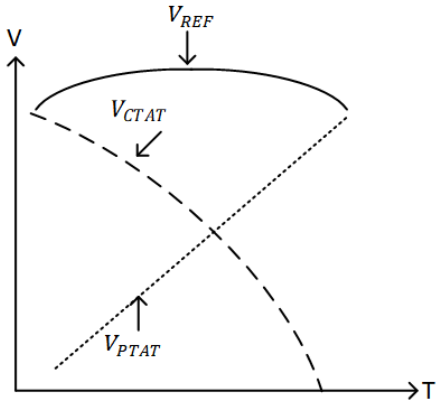

Figure 1.2: Voltage Reference Generation with Output Curvature

The traditional bandgap reference circuit has a V_{REF} of about 1.2V, which is close to the theoretical bandgap of silicon at 0K. As the channel length in modern processes has decreased, so too has the supply voltage of these devices. The supply voltage is made smaller to decrease power consumption and also because of the very narrow gate oxide,

which, with a high voltage applied at the gate, could cause a very high electric field which would damage the device. As a result of this supply voltage reduction, the traditional bandgap voltage of 1.2V is no longer practical, considering the nominal supply voltage for the 45nm process node is 1V. It is no longer possible to generate this 1.2V constant voltage reference at low supply voltages. Therefore, it is important for analog ICs at these narrow channel length nodes to have a stable voltage reference across temperatures. This new reference voltage should operate at a voltage less than 1V, so new methods must be implemented for this to occur.

This thesis outlines the design, layout, and test results for a simple voltage reference circuit which implements a sub-1V voltage reference. This circuit is unique in its generation of PTAT and CTAT voltages, and is notable for using the threshold voltage of a MOSFET for the CTAT voltage. By using the threshold voltage of a MOSFET, that means that even across process variation which varies the threshold voltage, then the reference voltage will be sufficient to bias any device biased by the reference voltage in strong inversion because the reference voltage will also scale accordingly. The equations to describe the circuit's operation are shown and derived using the first-order long-channel model MOSFET equations, neglecting channel-length modulation. These equations model approximate behavior, and simulations are done using the BSIM4 models for TSMC180nm which model higher-order behavior. Measurements were taken on five chips across temperature, and the reference voltage circuit has an average measured voltage variation of 4.62mV over a 80°C range, which makes it suitable for the prior mentioned 6-bit data converter which has an LSB of about 16mV. With this amount

of variation, it would not be suitable for applications in very high-resolution circuits, but for a 6-bit ADC, it would be very suitable.

This thesis also outlines the operation and simulation results for a variation of the simple sub-1V voltage reference circuit which replaces the BJTs with MOSFETs, which allows for a much smaller area for the circuit compared to the one with BJTs. This circuit is designed in the TSMC40nm process. The simulation results indicate less voltage variation across temperature and less chip-to-chip variation than the BJT implementation. These improvements, combined with the area reduction, mean it could have better performance than the prior circuit, although it needs to be taped out to verify performance in the actual silicon.

1.2 Organization of the Thesis

Chapter 2 details the traditional bandgap reference circuit implementation and theory. These circuits each have an output voltage of around 1.2V. Chapter 3 introduces some implementations to make a sub-1V bandgap circuit, and chapter 4 introduces the design of the proposed, simple voltage reference circuit. Chapter 5 shows the measured results obtained from silicon fabricated in the TSMC 0.18µ process. Chapter 6 details the variation of the circuit introduced in chapter 4 which replaces the BJTs with MOSFETs. Chapter 7 gives the closing remarks and conclusion.

Chapter 2: Traditional Bandgap Reference Circuits

Prior to the development of the bandgap reference circuit, temperature-compensated Zener diodes were used as reference voltages [1]. Zener diodes have multiple issues that make them undesirable for use as stable voltage references. In relation to today's narrow channel length devices, Zener diodes have breakdown voltages much greater than 1V, which makes them impractical as they would damage the devices. Additionally, Zener diodes are noisy and difficult to maintain a tight process control [1]. These were the motivations for the development of a bandgap reference circuit. This chapter will explore the bandgap reference circuit implementing a full value bandgap voltage around 1.2V, starting from the original one proposed by Widlar, and then into further considerations including CMOS implementations and start-up circuits for the bandgap circuit.

2.1 CTAT AND PTAT VOLTAGE GENERATION

This section details the traditional method of generating a CTAT and PTAT voltage. While there are many candidates for components and circuits with temperature coefficients, BJTs have historically been used because of their well-known behavior over temperature.

2.1.1 CTAT Generation

The standard equation for V_{BE} is shown below in equation (2.1).

$$V_{BE} = \frac{kT}{q}\ln\left(\frac{I_C}{I_S}\right) \tag{2.1}$$

At first glance, V_{BE} looks PTAT because of the kT/q term, which increases with temperature. This term, kT/q, is also known as the thermal voltage, written as V_T. However, it is important to consider the reverse saturation current term, I_S. I_S can be rewritten in the equation in terms of its temperature components, as below, where V_{G0} is the extrapolated bandgap voltage at 0K and is 1.205V [2]. Equation (2.2) can be simplified further and is shown in (2.3).

$$V_{BE} = \frac{kT}{q}\ln\left(\frac{I_C}{I_0}e^{V_{G0}/\left(\frac{kT}{q}\right)}\right) \tag{2.2}$$

$$V_{BE} = V_{G0} - \frac{kT}{q}\ln\left(\frac{I_0}{I_C}\right) \tag{2.3}$$

It can be shown that the change in V_{BE} over temperature is as follows.

$$\frac{dV_{BE}}{dT} = \frac{-k}{q}\ln\left(\frac{I_0}{I_C}\right) \tag{2.4}$$

Equating (2.4) with V_{BE} -V_{G0} from (2.3), the change in voltage over temperature is shown below.

$$\frac{dV_{BE}}{dT} = \frac{V_{BE}-V_{G0}}{T} \tag{2.5}$$

Ultimately, V_{BE} is CTAT due to the temperature factor in the denominator of equation (2.5), so as the temperature increases, V_{BE} will decrease.

2.1.2 PTAT Generation

There is not any need for extra circuitry to generate a CTAT voltage because the bipolar V_{BE} is intrinsically CTAT. On the other hand, a PTAT voltage requires a circuit to

generate it. Consider figure 2.1, where Q_1 and Q_2 are identical except for the current through them.

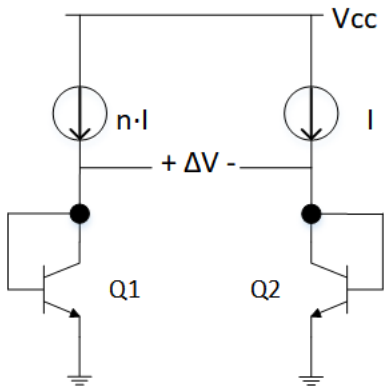

Figure 2.1: Generation of PTAT Voltage with Collector Current Scaling

Because of the diode connection from the collector to base of the BJTs, the differential voltage can be written as

$$\Delta V = V_{BE1} - V_{BE2} \tag{2.6}$$

Rewriting the base-emitter voltages in terms of the currents,

$$\Delta V = \frac{kT}{q}\ln\left(\frac{I_{c1}}{I_s}\right) - \frac{kT}{q}\ln\left(\frac{I_{c2}}{I_s}\right) = \frac{kT}{q}\ln\left(\frac{I_{c1}}{I_{c2}}\right) \tag{2.7}$$

Therefore, this ΔV itself is PTAT since the only temperature component is kT/q, so as temperature increases, the ΔV will increase. Another case for PTAT generation is if the collector currents of two transistors are the same, but the area of Q_2 is m times that of Q_1, as shown in figure 2.2.

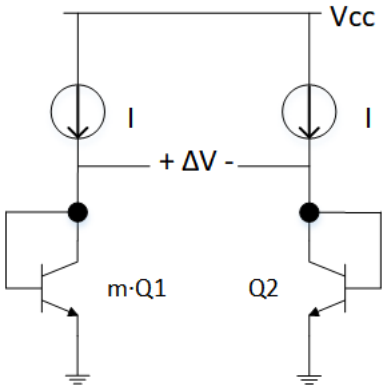

Figure 2.2: Generation of PTAT Voltage with Device Scaling

In this case, the difference in voltage is

$$\Delta V = \frac{kT}{q}\ln\left(\frac{I_{C1}}{I_{S1}}\right) - \frac{kT}{q}\ln\left(\frac{I_{C2}}{I_{S2}}\right) = \frac{kT}{q}\ln\left(\frac{I_{S2}}{I_{S1}}\right) \tag{2.8}$$

Thus, a PTAT voltage can also be generated just by changing the physical area of the BJT, which is directly related to the reverse saturation current of the BJT. These two methods, increasing the collector current of Q_1, relative to Q_2, and increasing the device area of Q_2, relative to Q_1, can be combined. For example, if $I_{C1} = n \cdot I_{C2}$ and $Q_2 = m \cdot Q_1$, then the PTAT voltage can be written

$$\Delta V = \frac{kT}{q}\ln\left(\frac{n*I_{C2}}{I_{S1}}\right) - \frac{kT}{q}\ln\left(\frac{I_{C2}}{m*I_{S1}}\right) = \frac{kT}{q}\ln(m*n) \tag{2.9}$$

Therefore, the PTAT voltage can be controlled by choosing the current through Q_1 and the sizing of Q_2.

2.2 Classic Widlar Bandgap Reference Circuit

In his 1971 work, "New Developments in IC Voltage Regulators", Bob Widlar introduced a simple circuit to implement a reference voltage that will vary little over a wide range of temperature [1]. This circuit is shown in Figure 2.3.

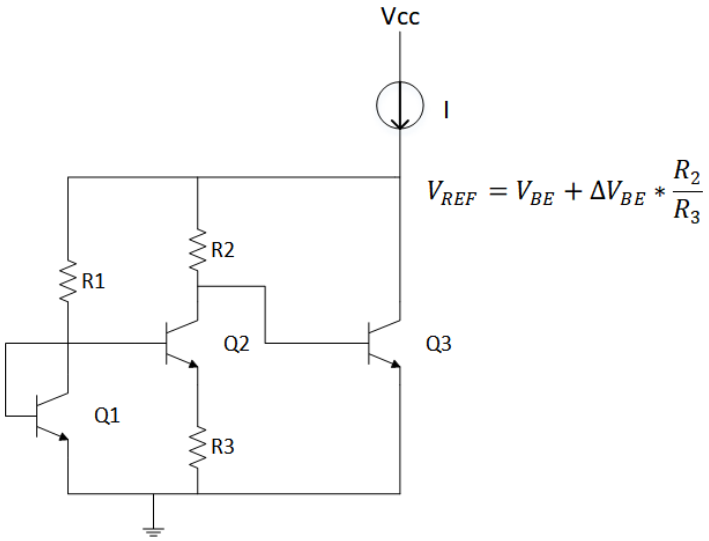

Figure 2.3: Original Bandgap Reference Simple Circuit Diagram

This circuit uses three bipolar junction transistors (BJTs) for implementing the PTAT and CTAT components. The PTAT component comes from Q_1 and Q_2. Writing Kirchhoff's Voltage Law (KVL) around Q_1, Q_2, and R_3, it is found that the voltage across R3 is

$$V_{R3} = V_{BE1} - V_{BE2} = \Delta V_{BE} \tag{2.10}$$

Assuming a large current gain for Q_2, which implies $\beta >> 1$, then the voltage across R2 is

$$V_{R2} = \frac{\Delta V_{BE}}{R_3} * R_2 \tag{2.11}$$

Next, an additional KVL is written considering Q_3, R_2, and the output voltage, V_{REF}. This equation shows

$$V_{REF} = V_{BE} + \Delta V_{BE} * \frac{R_2}{R_3} \tag{2.12}$$

Furthermore, the difference of the base-emitter voltages can be written in terms of their current densities, expressed as J_1 and J_2. The current density is the ratio of the collector current to the BJT reverse saturation current.

$$\Delta V_{BE} = \frac{kT}{q} * \ln(\frac{J_1}{J_2}) \tag{2.13}$$

Looking at only the temperature varying components of V_{REF}, then its equation becomes

$$V_{REF} = V_{BE} + \frac{kT}{q} * \ln\left(\frac{J_1}{J_2}\right) = V_{g0}\left(1 - \frac{T}{T_0}\right) + V_{BE0}(\frac{T}{T_0}) + \frac{kT}{q} * \ln\left(\frac{J_1}{J_2}\right) \tag{2.14}$$

The resistors, R_2 and R_3 are not included because their temperature coefficients will cancel out. Ultimately, it is shown that, by setting the derivative of V_{REF} with respect to temperature equal to zero, the output will be totally temperature compensated provided

$$V_{g0} = V_{BE0} + \frac{kT_0}{q} * \ln\left(\frac{J_1}{J_2}\right) \tag{2.15}$$

As long as the sum of the initial base-emitter voltage and the PTAT term are equal to the bandgap voltage of silicon at 0K, then the reference voltage provided will be stable over temperature. In practice, due to higher order effects not considered in the derivation, the output voltage will be made slightly higher than 1.205V [1].

In terms of actual values, a typical temperature variation of a V_{BE} is -1.5mV/K at room temperature [3]. This means the PTAT portion should have a temperature

coefficient of +1.5mV/K. In order for this to be achieved, then $\frac{kT}{q} * \ln(\frac{J_1}{J_2})$, which can be re-expressed in units of mV/K as $.087 * \ln(\frac{J_1}{J_2})$mV/K, needs to be +1.5mV/K. For this to occur, the current density ratio would need to be $3.07E^7$ [3]. Because the current density comes from scaling currents and device sizing, it is very impractical to achieve this kind of current density ratio in an IC. This shows the need of having a scaling factor for the PTAT voltage so that the current density ratio can be reduced. From equation (2.12), the PTAT voltage in the circuit is indeed scaled by the ratio of R_2 and R_3. By choosing their ratio to be 10, for example, the requirement for the current density is reduced to be 5.61, a significant improvement and very realizable in a circuit.

One major assumption made in the derivation of the bandgap reference voltage was that the beta factor, which is the ratio of the collector current to the base current, for each BJT was large, so each parallel current branch was not negatively affected by the others. The beta for BJTs do not match well and have a very wide process distribution [4]. This makes it practically difficult to achieve a temperature independent reference voltage. Additionally, there is no feedback mechanism in this circuit, so any variation will not be corrected by negative feedback. The PSRR of this circuit is also an issue because an change in the supply will change the current flow into the three branches. If the current source was ideal, this would not be an issue, but because it is not ideal, any change in the power supply will induce a change in the current, which is directly related to the base-emitter voltages.

2.3 CMOS Implementation of Bandgap Reference Circuit

Once CMOS devices started to grow in popularity, there were attempts to implement the bandgap circuit using only CMOS devices. The circuits presented in this section still use BJTs for generating the temperature coefficients, but also use MOSFETs for creating a stable reference voltage equal to the bandgap voltage while temperature varies.

2.3.1 Original CMOS Bandgap Circuit

In this circuit, the final output is very similar to Widlar's original circuit, although the circuit is quite different. This circuit is shown in figure 2.4 [5].

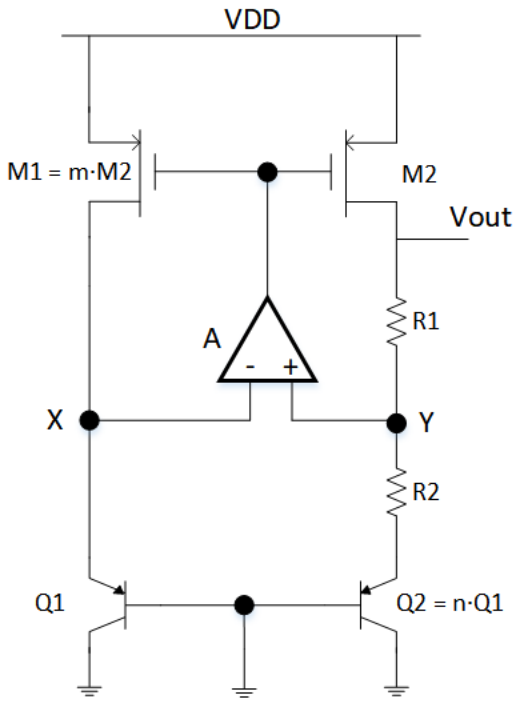

Figure 2.4: CMOS Implementation of Bandgap Reference

The bipolars in this design are pnp devices and there is no dependence on beta for this circuit since the pnp devices are vertical and do not share current. The key of this design is the operational amplifier which is used to force nodes X and Y to be equivalent, so that ΔV_{BE} can be sensed across the resistor R_2. Assuming an ideal op amp, with infinite gain and no offset error, the output voltage can be found via KVL, where V_{R2} and V_{R1} are voltages across R_2 and R_1, respectively.

$$V_{out} = V_{BE2} + V_{R2} + V_{R1} \tag{2.16}$$

Recognizing that the op amp will make the voltage at node Y equal to V_{BE1}, it is found

$$V_{out} = V_{BE2} + \Delta V_{BE} + \Delta V_{BE}\frac{R_1}{R_2} = V_{out} = V_{BE2} + \Delta V_{BE}\left(1 + \frac{R_1}{R_3}\right) \tag{2.17}$$

This expression comes from the fact that the current through M_2 and Q_2 is $\Delta V_{BE}/R_2$. If Q_2 is n times Q_1, and then M_1 is m times M_2, then once again, we can write that

$$\Delta V_{BE} = \frac{kT}{q}\ln(m * n) \tag{2.18}$$

Plugging (2.18) into (2.17) yields

$$V_{out} = V_{BE2} + \frac{kT}{q}\ln(m * n)\left(1 + \frac{R_1}{R_2}\right) \tag{2.19}$$

This the same as the original bandgap circuit, except for an additional factor of adding 1 to the resistor ratio R_1/R_2. An advantage of this circuit is that it requires only two BJTs instead of the three BJTs required by the original bandgap circuit introduced by Widlar.

2.3.2 Issues of Simple CMOS Bandgap Reference

Ideally, a bandgap reference circuit will also have a very high power-supply rejection ratio (PSRR). In this circuit, the weak link for PSRR is the op amp itself. The circuit will be easily affected by a lower PSRR. Additionally, the prior analysis assumed an infinite op amp gain. For shorter channel length process which have lower supply voltages, the intrinsic gain of MOSFETs decreases, so it is harder to achieve a high gain. Additionally, the offset error of the op amp will be amplified by $(1 + R_1/R_2)$ [5]. This means the designer either needs to guarantee a very small offset error on the op amp or make a high (m*n) factor, which is the product of the area difference between the two bipolar devices and the two MOSFETs. Lastly, this circuit has a potential issue when the circuit starts. If the output of the op amp is high, then the two pMOS devices, M_1 and M_2, will not be able to turn on, so it will be a steady state of operation where there is no current flow in the two branches, which means that there will be no base-emitter voltage to generate the bandgap voltage. The solution for this is addressed in the next section.

2.3.3 Start-up Circuit for Simple CMOS Bandgap Reference

As mentioned in section 2.3.2, there exists a stable operating point of this circuit in which there will be no current flow [3]. In order to correct this, extra components are added that will help the circuit to start up into the correct stable operating point, and then turn off once that point is reached. One implementation is shown in figure 2.5 [3].

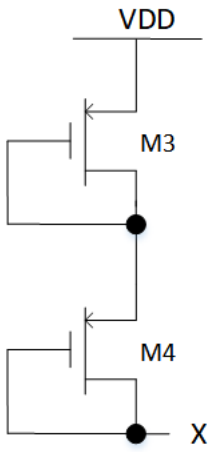

Figure 2.5: Bandgap Reference with Start-up branch

The function of this circuit is to create an initial imbalance between the two inputs of the op amp. This will force the output node low, which will then allow the top pMOS devices to turn on. Then, the circuit will move into the stable operating point at which node X and node Y in figure 2.4 will be equal to the base-emitter voltage of Q_1. Once this point is reached, then the start-up branch, M_3 and M_4, will turn off and the bandgap reference circuit operates as normal. As the circuit moves into its stable operating point, more current will be sourced from the top pMOS device, M_1, instead of M_3 and M_4. Once the circuit reaches a stable operating point, the current from M_1 will be equal to the current in Q_1. At this point, M_3 and M_4 will turn off.

Chapter 3: Developments in Sub-1V Bandgap Reference Circuits

The circuits presented in chapter 2 all dealt with reference voltages at the full bandgap voltage of approximately 1.2V. These circuits work well with supply voltages high enough to generate a 1.2V reference voltage, but once the supply becomes lower than 1.2V, it is not possible to generate this voltage. The supply voltages in ICs have indeed been shrinking due to smaller and smaller feature sizes of transistors. Even though a lot of analog components will use sizes much greater than the minimum channel length, the gate oxide is still reduced for the process, which means a smaller gate voltage is needed. Smaller channel lengths are being used because of the great advantages they provide in digital circuits, which will use the minimum channel length, so analog circuits must be made to work while using a small voltage supply. It is not possible to generate a full 1.2V bandgap reference voltage with a supply voltage of 1V, so a sub-1V bandgap must be designed. This chapter will detail previous designs done to implement a sub-1V bandgap which can be used with smaller process nodes.

3.1 CMOS Implementation with Resistor Divider

This design is an extension of the basic CMOS implementation presented in section 2.3 of this thesis. The basic idea is to divide the standard 1.2V bandgap voltage using a resistor divider [6]. The method for this is to take a PTAT current, generate a CTAT voltage, and use a voltage divider to add the PTAT and CTAT components while dividing the voltage to a lower value. This circuit is shown in figure 3.1.

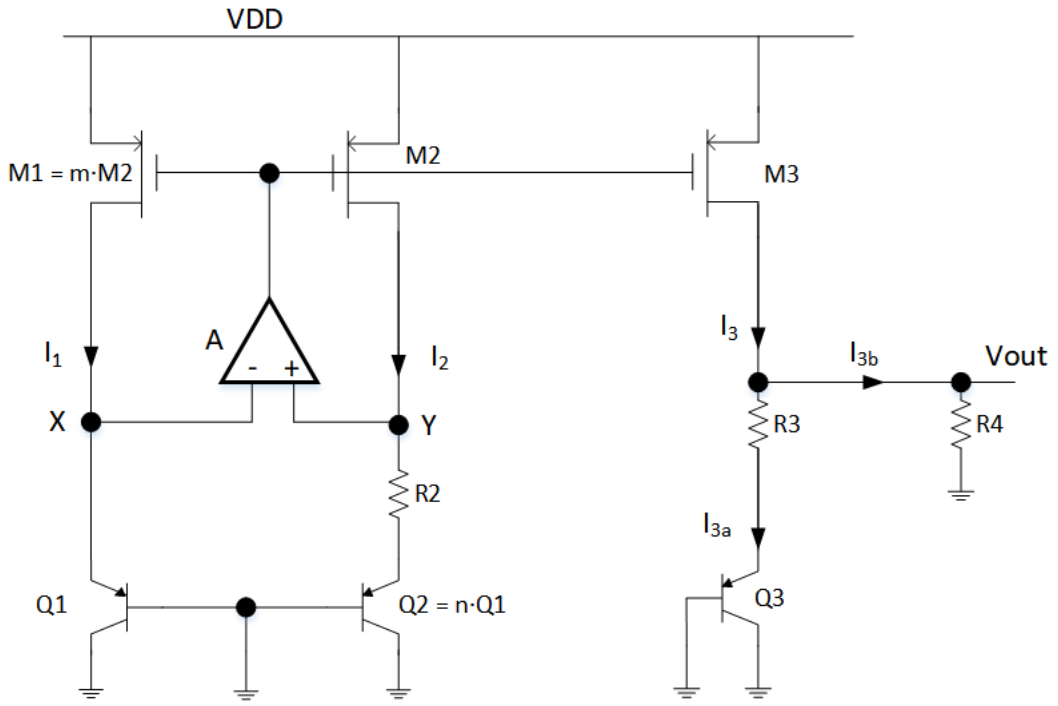

Figure 3.1: CMOS Bandgap Circuit with Voltage Divider for Sub-1V Reference Voltage

The left side of the circuit consisting of M_1, M_2, R_2, Q_1, Q_2, and A, is the same as presented in section 2.2, except it removes the resistor between the positive terminal of the op amp and the drain of M_2. The reason for not using that resistor is that now the purpose of the left part of the circuit is to generate a PTAT current and not a temperature-compensated output reference voltage. Assuming an ideal op amp with infinite gain, then the voltage at node X and Y are the same, which is equal to V_{BE1}. Therefore, the current, I_2, through Q_2 and M_2, is indeed PTAT and written as

$$I_2 = \frac{V_{BE1} - V_{BE2}}{R_2} = \frac{\Delta V_{BE}}{R_2} \tag{3.1}$$

Now, assuming the two pMOS transistors, M_2, and M_3, are well-matched, then the currents, I_2, and I_3, should all be equal.

$$I_2 = I_3 \tag{3.2}$$

Writing KCL at the output voltage node,

$$I_3 = I_{3a} + I_{3b} \tag{3.3}$$

Using Ohm's law and plugging in from (3.1) and (3.2), it is found

$$\frac{\Delta V_{BE}}{R_2} = \frac{V_{out} - V_{BE3}}{R_3} + \frac{V_{out}}{R_4} \tag{3.4}$$

Through algebraic simplification and solving for the output voltage,

$$V_{out} = \frac{R_4}{R_3 + R_4}\left[V_{BE3} + \frac{R_3}{R_2} * \Delta V_{BE}\right] \tag{3.5}$$

Therefore, by appropriate choice of R_3 and R_4, a lower voltage reference value can be obtained, while still preserving the temperature compensation by adding a PTAT and CTAT voltage.

3.1.1 Issues of CMOS Reference Circuit with Resistor Divider

The circuit just introduced in section 3.1 has a few notable issues. The reason for needing to divide the bandgap voltage is that the supply voltage is shrinking. As the supply voltage decreases, it is harder to design the op amp because the swing is reduced. Additionally, it will be more difficult to obtain the high gain required to make the derived equations accurate. Furthermore, this design has an area overhead of adding an extra BJT and resistor. Another key assumption made in the analysis in section 3.1 was there the pMOS devices are very well-matched in order to make the PTAT currents match very accurately.

3.2 Sub-1V Bandgap Reference Circuit Using Current Summation

This sub-1V bandgap reference circuit also uses an op amp and two branches with different sized BJTs for generating a PTAT voltage. The BJTs here are implemented as substrate pnp BJTs, and are often drawn as diodes because both the collector and base are connected to ground [see the Appendix for details on substrate pnp devices]. The fundamental difference of this circuit is that it uses the summation of CTAT and PTAT currents instead of voltages [7]. The output voltage is produced by mirroring that current and then passing it through a resistor. This circuit is shown in figure 3.2.

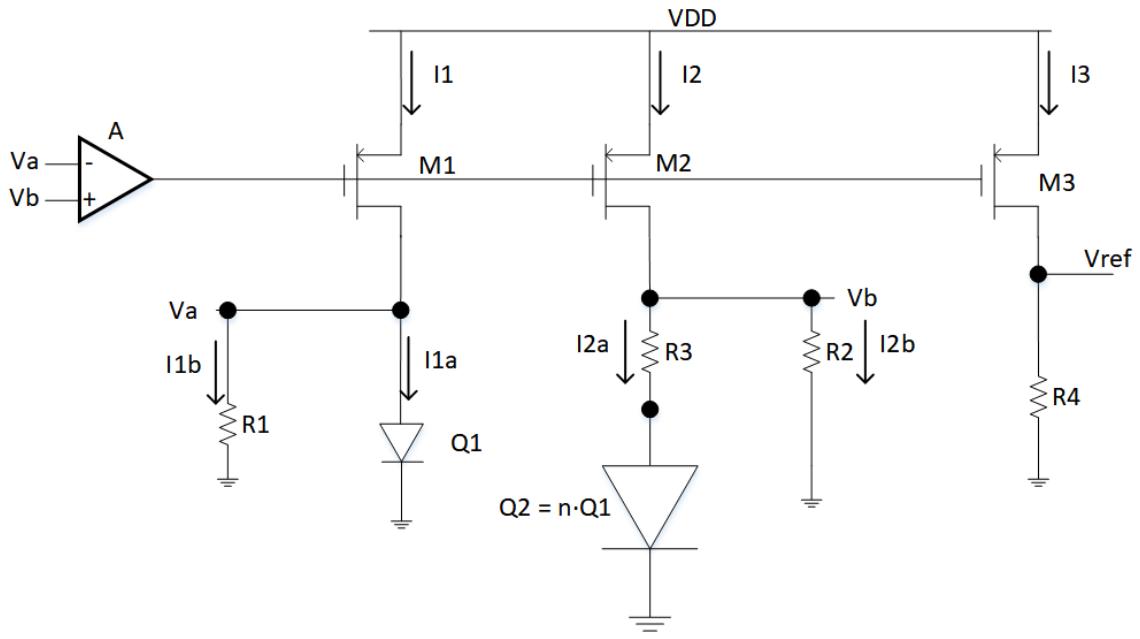

Figure 3.2: Schematic of Sub-1V Bandgap Circuit using Current Summation

In the circuit in figure 3.2, R_1 is chosen to be the same as R_2 and the three pMOS devices are designed the same and are assumed to have very good matching, so that

$$I_1 = I_2 = I_3 \tag{3.6}$$

Because of the ideal op amp, the voltage at node a and node b will be forced to be the same, and the voltage at node a is V_{BE1}. Therefore,

$$I_{2a} = \frac{V_{BE1} - V_{BE2}}{R_3} = \frac{\Delta V_{BE}}{R_3} \tag{3.7}$$

$$I_{2b} = \frac{V_{BE1}}{R_2}$$

$$I_2 = \frac{\Delta V_{BE}}{R_3} + \frac{V_{BE1}}{R_2}$$

Then, that current will be the same that goes through M_3 and R_4, so the output voltage is written

$$V_{REF} = \frac{R_4}{R_3} * \Delta V_{BE} + \frac{R_4}{R_2} * V_{BE1} \tag{3.8}$$

By appropriate choice of the resistors R_4,R_3, and R_2, the output can be both temperature compensated and less than 1V. In fact, the supply voltage can theoretically be as low as the base-emitter voltage of the BJT provided the reference voltage is less than that base-emitter voltage [7].

A key advantage of this circuit in terms of mismatch, is that the reference voltage depends only on the ratio of resistors, not the value of resistors. Resistors track each other across process variation, so their mismatch will be absorbed by taking the ratio of the resistance values. This circuit is an improvement on the one in section 3.1 because it requires one less BJT to implement a sub-1V reference voltage.

The disadvantages of this circuit are similar to the prior circuit in that it also requires and extra resistor and current branch, and also heavily relies on the op amp to make node a and node b equivalent.

Chapter 4: Design of Simple Voltage Reference Circuit

This chapter details the design of a new sub-1V voltage reference circuit, which is the core of this thesis. This circuit has a novel way of generating the PTAT component, and instead of using the base-emitter voltage of a BJT for the CTAT voltage, it instead uses the threshold voltage of a MOSFET which is also CTAT. This circuit requires no resistors to save on area and no start-up circuit because a stable operating point is guaranteed by design. The core of this circuit generates a PTAT current, that can be mirrored and passed into a diode-connected device, which will produce a voltage reference stable across temperature.

4.1 Principle of Operation

As with any voltage reference circuit which is to be stable across a temperature range, the fundamental principle of operation involves adding a CTAT voltage with a scaled PTAT voltage to achieve a reference voltage with minimal variation over temperature. This is expressed in the equation below, where V_{BE1} is the CTAT voltage, V_{BE1} - V_{BE2} is the PTAT voltage, and λ is the scaling factor.

$$V_{REF} = V_{BE1} + \lambda * (V_{BE1} - V_{BE2}) \tag{4.1}$$

Here, V_{BE1} and V_{BE2} are voltages across two diode-connected substrate pnp transistors. j_1 and j_2 are current densities of V_{BE1} and V_{BE2}. These transistors are biased at a fixed current density ratio $\gamma = (j_1/j_2)$. As shown in chapter 2, the difference of two base-emitter voltages is indeed PTAT. Likewise, V_{BE1} will decrease somewhat linearly with

temperature when the diode is biased with constant current, and is an example of a CTAT voltage. Another example of a CTAT voltage is the threshold voltage, V_{th}, of a MOSFET. It has been shown in [8] that the change of threshold voltage with respect to change in temperature can be written as

$$\frac{dV_{th}}{dT} = \frac{d\Phi_F}{dT} + \left[1 + \frac{q}{C_{ox}} * \sqrt{\frac{\varepsilon_{Si} N_a}{kT * \ln\left(\frac{N_a}{n_i}\right)}}\right] \tag{4.2}$$

In equation (4.2), Φ_F is the Fermi Level, q is the charge of an electron, ε_{Si} is the dielectric constant of silicon, N_a is the dopant concentration, k is Boltzmann's constant, C_{ox} is the capacitance of the gate oxide, T is temperature in Kelvins, and n_i is the intrinsic carrier concentration. Furthermore, as in [8], the change in Fermi Level over change in temperature can be written as

$$\frac{d\Phi_F}{dT} = 8.63 * 10^{-5} * \left[\ln(N_a) - 38.2 - \frac{3}{2}\{1 + \ln(T)\}\right] \tag{4.3}$$

Both the first and second terms on the right side of equation (4.2) will decrease with temperature, so the threshold voltage is indeed CTAT. It has been shown that a typical temperature coefficient of the threshold voltage in nMOS, which is the device used in this design, is -2mV/°C [9]. In this design, V_{th} is used instead of V_{BE1}. This is done for several reasons such as it avoids the need of a resistive voltage divider to reduce V_{BE}, which is what is done in many sub-1V designs. Additionally, most short-channel technologies provide a threshold voltage between 0.4V to 0.5V, which is the lower value needed for a reference voltage between 0.6V and 0.8V, which is achievable with a 1V supply. Also, the fact that the reference voltage is based on the threshold voltage means that transistors

whose gates are driven by the reference voltage will operate in strong inversion. As the threshold voltage is varied over process corners, then that means the reference voltage will also be varied across process corners, so the devices can still be biased in strong inversion. For a non-V_{th}-based reference voltage, if the threshold voltage is increased, and the reference voltage does not scale accordingly, then it is possible that that device will not be turned on at all, so no current will flow in that particular branch. This is the importance of having a V_{th}-based reference voltage in MOS ICs.

Let the base-emitter voltage differential be represented with the symbol V_d.

$$V_{REF} = V_{th} + \lambda * (V_d) \tag{4.4}$$

Therefore, the three operations needed to obtain the reference voltage are subtraction to obtain V_d, multiplication by λ, and then addition of the CTAT threshold voltage. To implement these operations, there are many different circuit implementations for applications with varying levels of precision. Operational amplifiers with very low offset, drift, and noise may be used to obtain very high precision references for high precision data converters. In this thesis, the circuit is suitable for moderate precision applications such as a 6-bit ADC, such as those used in high-speed data communication systems that require sampling rates above 1GHz.

The gain, λ, needed to implement the full bandgap value of 1.2V is approximately 5, if V_{BE1} - V_{BE2} is approximately 100mV, which can be achieved by a current density of $j_1/j_2 = 100$. For a fractional bandgap reference voltage, it can be seen by scaling both sides of equation (4.1), that the gain required will be proportionately less.

Equation (4.1) is written assuming that the pn junctions are biased with a constant bias current. In this design, the key idea is to establish a constant voltage reference across a diode-connected MOS device. To generate this voltage reference, a temperature-dependent bias current needs to be generated that will give a constant reference voltage between 600mV and 800mV depending on the threshold value of the MOS device. The generation of this bias current with appropriate temperature-dependence by a simple circuit is the key innovation of this design.

4.2 Design Of Voltage Reference Circuit

The voltage reference circuit is shown in figure 4.1. The core of the circuit, which generates the PTAT component, is shown in the dotted lines. The branch outside of the dotted line is used for passing the PTAT current generated by the core circuit through a diode-connected device to generate the reference voltage output.

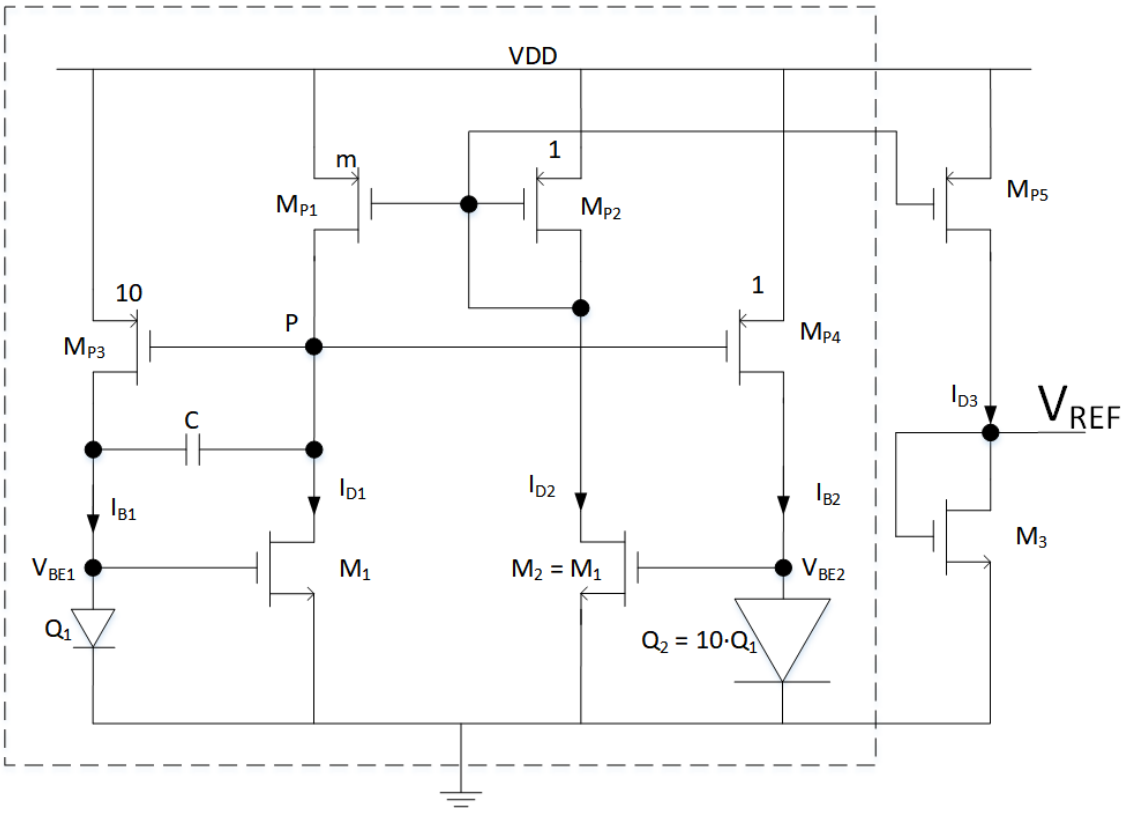

Figure 4.1: Circuit Schematic of Proposed Voltage Reference Circuit

In the circuit, Q_1 and Q_2 are emitter-base junction diodes of substrate pnp transistors biased by currents I_{B1} = 10I and I_{B2}= I, respectively. These bias currents are obtained by a pair of pMOS transistors, M_{P3} and M_{P4}, which are sized in a ratio of *10:1*. By forcing Q_2= $10Q_1$, then the current density, j_1/j_2, where j_1 is the current density of Q_1, and j_2 is the current density of Q_2, is *100:1* and is built into the circuit purely by sizing of devices. Thus, the base-emitter voltages can be written as

$$V_{BE1} = V_T \ln\left(\frac{10I}{I_S}\right) \tag{4.5}$$

$$V_{BE2} = V_T \ln\left(\frac{I}{10I_S}\right)$$

The difference of these base-emitter voltages is written

$$V_{BE1} - V_{BE2} = V_T \ln\left(\frac{\frac{10I}{I_S}}{\frac{I}{10I_S}}\right) = V_T \ln(100) \tag{4.6}$$

As shown before, this difference in base-emitter voltages is indeed a PTAT voltage. The magnitude of this PTAT voltage is indeed $V_T\ln(100)$ which is simple to implement, and it is also difficult to obtain a much larger value due to the nature of the logarithmic function. Therefore, as in the traditional bandgap implementations, some gain is required for the PTAT voltage so that it relaxes the requirement to generate a larger value solely from the logarithmic function.

Re-writing equation (4.1) as

$$V_{REF} = \frac{1}{2}(V_{BE1} + V_{BE2}) + \lambda' * (V_{BE1} - V_{BE2}) \tag{4.7}$$

From this equation, it is noted that both the differential mode and common mode gains will be needed. This suggests the need for an amplifier topology that will have a common mode gain of unity and differential mode gain of λ'. This can be implemented by just removing the current tail bias of the standard differential pair. Thus, the base-emitter voltages, V_{BE1} and V_{BE2} are applied as gate-source voltages to identical nMOS transistors, M_1 and M_2, with grounded sources. The input devices have active load devices, M_{p1} and M_{p2}. M_{p1} has a size of *m* and M_{p2} is a unit size of 1. This circuit configuration looks like the familiar differential pair used for amplifying a differential voltage, v_d, but it does not fully reject the common mode; instead it has a finite common mode gain. Let the drain currents of the branches be I_{D1} of M_1 and I_{D2} of M_2. If long channel devices are chosen and operated in the strong inversion region, then an ideal

square-law model can be assumed for the transistors. Ignoring any channel length modulation effects, the drain current of M_1 is written as

$$I_{D1} = k * (V_{BE1} - V_{th})^2 \tag{4.8}$$

In this equation,V_{th} is the threshold voltage, and *k* has its usual meaning as the transconductance parameter which is

$$k = \frac{1}{2}\mu_n C_{ox}\frac{W}{L} \tag{4.9}$$

In (4.9), μ_n is the electron mobility, C_{ox} is the gate capacitance, W is the width of the device, and L is the length of the channel. Similarly, I_{D2} can be written

$$I_{D2} = k * (V_{BE2} - V_{th})^2 \tag{4.10}$$

This current is inverted by the using the p-channel current mirror consisting of M_{P1} and M_{P2}, with a gain of *m*, and made to fight with I_{D1} at node *P.* The current $I_{D1} – I_{D2}$ is written as

$$I_{D1} - I_{D2} = k * (V_{BE1} - V_{th})^2 \ - \ \mathrm{m} * k * (V_{BE2} - V_{th})^2 \tag{4.11}$$

Since v_d is positive due to V_{BE1} being larger, I_{D1} is greater than I_{D2}, the voltage at node P will drop, which causes M_{P3} and M_{P4} to turn on, which establishes quiescent currents in Q_1 and Q_2. For m=1, the current I_{D1} will always be greater than I_{D2} since v_d remains constant. Thus, the bias currents in the diodes will keep on increasing to keep their ratio constant. Using the algebraic identity x^2-y^2= (x+y)(x-y), then the right hand side of (4.11) can be rewritten as

$$I_{D1} - I_{D2} = 2kv_d\left[\frac{1}{2}(V_{BE1} + V_{BE2}) - V_{th}\right] \tag{4.12}$$

For m=1, the common mode voltage will continue to increase until the devices enter the triode mode and then invalidate the assumptions. This means that a value of *m>1* should be chosen. This will be shown below.

Note that there are two feedback loops in the circuit. One contains M_{P3}, Q_1, and M_1, and the other contains M_{P4}, Q_2, M_2, as well as the current mirror (M_{P2}, M_{P1}). The first loop provides positive feedback, and the second provides negative feedback. The positive feedback results in a negative conductance from node P to ground and the negative feedback has a positive conductance from node P to ground. Because the negative feedback experiences the gain from the current mirror, the negative feedback will be larger than that of positive feedback. In order to suppress any oscillations in the positive feedback path, a small capacitor of 1pF is connected between the gate and drain of M_1. Thus, the circuit will settle in a stable state at all temperatures when $I_{D1}=mI_{D2}$. This gives

$$k * (V_{BE1} - V_{th})^2 = \text{m} * k * (V_{BE2} - V_{th})^2 \tag{4.13}$$

Taking the square root and cancelling *k*,

$$(V_{BE1} - V_{th}) = \sqrt{m} * (V_{BE2} - V_{th}) \tag{4.14}$$

Subtracting (V_{BE2}-V_{th}) from both sides,

$$v_d = (V_{BE2} - V_{th})\left[\sqrt{m} - 1\right] \tag{4.15}$$

Then, representing $1/\left[\sqrt{m} - 1\right]$ as (λ-1),

$$V_{BE2} = V_{th} + v_d(\lambda - 1) \tag{4.16}$$

Then, writing in terms of V_{BE1},

$$V_{BE1} = V_{th} + \lambda v_d \tag{4.17}$$

From the equations for the two base-emitter voltages, which are the gate-source voltages of M_1 and M_2, the excess voltages of M_1 and M_2 will be in the ratio of $\lambda/(\lambda-1)$. By substituting these two equations into the drain current equation for M_1 and M_2,

$$I_{D1} = k * (V_{th} + \lambda v_d - V_{th})^2 = k * \lambda^2 {v_d}^2$$

$$I_{D2} = k * (V_{th} + v_d(\lambda - 1) - V_{th})^2 = k * (\lambda - 1)^2 {v_d}^2 \quad (4.18)$$

Using equations (4.16) and (4.17), the common mode input voltage can be written as

$$V_{cm} = \frac{V_{BE1}+V_{BE2}}{2} = V_{th} + v_d(\lambda - \frac{1}{2}) \quad (4.19)$$

Next, the gain λ for the PTAT voltage, v_d, as a function of the gain *m* of the current mirror is written as

$$\lambda = \frac{\sqrt{m}}{\sqrt{m}-1} \quad (4.20)$$

A plot of this function is shown in figure 4.2.

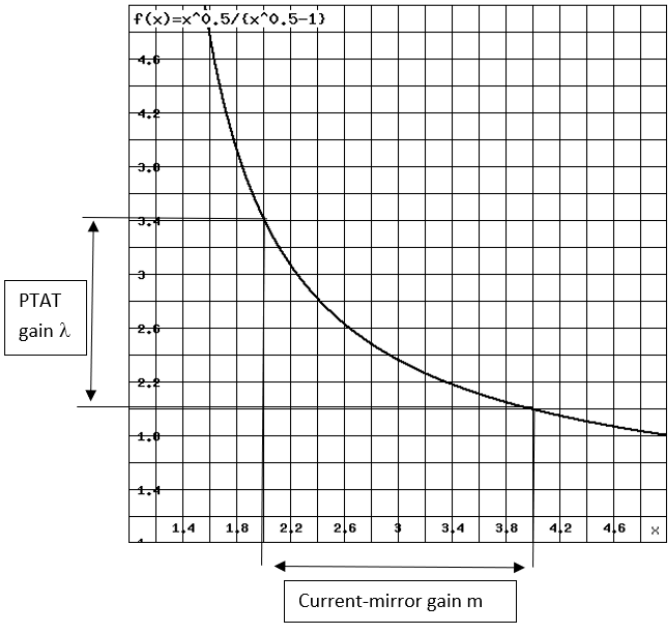

Figure 4.2: Plot of PTAT gain as a function of Current-mirror gain

For current mirror gains of 1<m<2, very high PTAT gains can be achieved. However, as mentioned earlier, a PTAT gain less than 5 can be used to implement a sub-1V reference voltage. Modest values of PTAT gains needed for a low-voltage reference are between 2 and 3.4, which correspond to a current-mirror gain between 2 and 4. The current-mirror gain is set just by the sizing of the p-channel devices, which is translinear, as well as process and temperature insensitive, so a PTAT gain of $2<\lambda<3.4$ is robust and can be easily trimmed. Now that the PTAT gain and current-mirror gain have been related, note that as in equations (4.16) and (4.17), the diode voltages V_{BE1} and V_{BE2} can be made constant by choosing m appropriately.

Now that the PTAT voltage has been generated from the core circuit, and I_{D1} and I_{D2} are both related to the PTAT voltage, I_{D2} is mirrored to M_{P5} by passing the gate voltage of M_{P2} to the gate of M_{P5}. The sub-circuit of this is shown in figure 4.3.

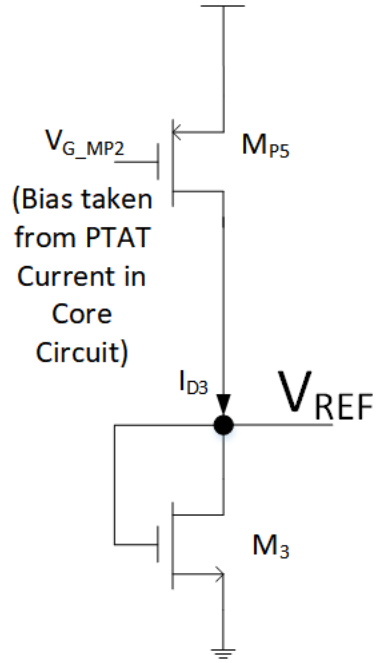

Figure 4.3: Sub-circuit for CTAT Generation of Proposed Voltage Reference Circuit

By using the long-channel model, the current relationship between I_{D2} and the current through M_{P5}, I_{D3}, is written as

$$I_{D3} = \frac{k_{p5}}{k_{p2}} * I_{D2} \tag{4.21}$$

In (4.21), k_{p5} and k_{p2} are the transconductance parameters for M_{P5} and M_{P2}, respectively. If the devices are perfectly matched, then the current through M_{P5} be an exact copy of I_{D2}.

The reference voltage is taken as the voltage across the diode-connected nMOS M_3. Because it is diode-connected,

$$V_{REF} = V_{GS} = V_{DS} \tag{4.22}$$

Therefore, as long as the reference voltage is greater than the threshold voltage, which will always be the case provided the supply voltage is greater than an over-drive voltage for the pMOS added with a threshold voltage for the nMOS, then M_3 will always be in saturation. Using the long-channel model,

$$I_{D3} = k_3 * (V_{GS} - V_{TH})^2 = k_3 * (V_{REF} - V_{TH})^2 \tag{4.23}$$

Solving for the reference voltage, where k_3 is the transconductance parameter for the diode-connected nMOS,

$$V_{REF} = \sqrt{\frac{I_{D3}}{k_3}} + V_{TH} \tag{4.24}$$

Using equation (4.21) to substitute for I_{DP3}, and equation (4.18) to then substitute for I_{D2}, and cancelling the transconductance parameter for M_{P2}, the reference voltage is written as

$$V_{REF} = v_d(\lambda - 1)\sqrt{\frac{k_{p5}}{k_3}} + V_{TH} \tag{4.25}$$

Therefore, the reference voltage does indeed consist of the sum of a scaled PTAT term and a CTAT term. Comparing this to equation (4.4), the circuit does implement the desired function: generation of PTAT voltage, scaling by a factor, and then adding it to a CTAT term, the threshold voltage of the diode-connected nMOS device. As seen from equation (4.25), because the PTAT gain seen at the output is actually (λ-1), so the PTAT gain needs to be increased to reach a reference voltage value between 600mV and

800mV. The gain can be increased by choosing a larger device width for M_{P5} to counteract the decrement of the designed PTAT gain, λ, which is determined by the width of M_{P1}.

4.3 Trimming Method for Simple Voltage Reference Circuit

Trimming in an IC design is used in post-silicon testing to adjust the value of some critical circuit blocks to meet the required specification. In many bandgap circuits, resistor values or MOS devices can be trimmed to correct for temperature variation or the magnitude of the reference value. By looking at the output equation for the reference voltage in this design, as in equation (4.25), there are a few process dependent terms. Because the threshold voltage is process dependent, the CTAT component varies with process, so the *k* value of M_{P5} or M_3 can be trimmed to obtain the constant reference voltage V_{REF}. It is not possible to trim the threshold voltage for the nMOS device, so only the PTAT component is trimmable.

The method for "trimming" in this design is to make use of the fact that each voltage reference circuit is small, so it is not too costly to just add multiple copies of the circuit onto the chip. The "trim" can be done by adjusting the width of the diode-connected nMOS device, which affects the k_3 term in equation (4.25). If the width increases, then V_{REF} should decrease, and if it decreases, V_{REF} should increase. The range over which the output nMOS device's width is changed is determined by simulation over process corner. The goal is that by giving enough variation of the PTAT component, a particular reference circuit will be better suited to serve as the reference voltage circuit to be used

for that whole lot of chips. Once the IC is fabricated, it must be at a particular process corner, so by giving enough variation in the PTAT component, at least one of the voltage reference circuits fabricated should be suitable to be the voltage reference for that chip. In a final application, this would be done by using a permanent trim that will make a permanent connection to the chosen voltage reference, so that it does not have to be re-selected each time the circuit is turned on.

4.4 Design Considerations

For the design of this voltage reference circuit, it is important to use long-channel devices so that the equations derived earlier will be valid. The equation derivations used the long-channel, square-law model and neglected channel-length modulation, so the channel length chosen for this design was 2µm, much higher than the process minimum of 180nm. This means that even though there will still be process mismatch, a nanometer mismatch on a micron scale matters much less than on a nano-scale. The unit width of the pMOS devices is 1.44µm, and that of nMOS is 220nm. The widths were chosen to be small because the device area was already made large by having a long channel-length. The unit size for the substrate pnp BJT is 2µm for both length and width.

In this chip tape-out, there were 8 voltage reference circuits designed in TSMC 180nm. Eight voltage reference circuits are designed in TSMC 180nm. The difference between each of them is that the output diode-connected nMOS device has varying widths in 20nm steps, so the devices range from 360nm to 500nm. The goal is to provide

enough variation so that some of the references will be good across the temperature range for every process corner.

The size of M_{P5} is determined via simulation and using the fact that it is necessary for it to have a higher unit width than M_{P2} to compensate for the PTAT gain being (λ-1). It is found that for a multiplier of 5 (M_{P5} having 5 unit devices in parallel), the reference voltage output is always greater than 600mV, less than 800mV, and has a reasonable variation across temperature.

Because the number of pins are limited on the chip, which is in a A-QFP44 package, an analog test multiplexer (mux) is used to select the reference circuit by sending in a digital code. The digital code needed for each circuit is shown in table 4.1.

Voltage Reference Circuit Selected	Digital Input Code
V_{REF0}	0000 0011 0000 0000
V_{REF1}	0000 0011 1000 0000
V_{REF2}	0000 0011 0100 0000
V_{REF3}	0000 0011 1100 0000
V_{REF4}	0000 0011 0010 0000
V_{REF5}	0000 0011 1010 0000
V_{REF6}	0000 0011 0110 0000
V_{REF7}	0000 0011 1110 0000

Table 4.1: Digital Code for Selecting Voltage Reference Circuit from Analog Test Mux

4.5 Simulation Results

Simulation was done using Cadence Virtuoso with the TSMC180 process design kit (pdk). The simulation was done at the top, chip-level using transient simulation. Each circuit is selected by applying a digital input stimulus. Time is allocated for the analog muxing to occur and then also for the output of the test mux to settle at its final value.

Each circuit was simulated from 0°C to 80°C in steps of 10°C , and the results are compiled in the graphs below. Simulation was also done over the five typical process corners, typical (tt), fast-fast (ff), fast-slow (fs), slow-fast (sf), and slow-slow (ss). The models for these are defined by the foundry, TSMC. Together, the process and temperature corners provide data for how the circuit should perform. The first graph seen in figure 4.4 shows the magnitude of the reference voltage taken at 30°C for each process corner and each reference voltage circuit.

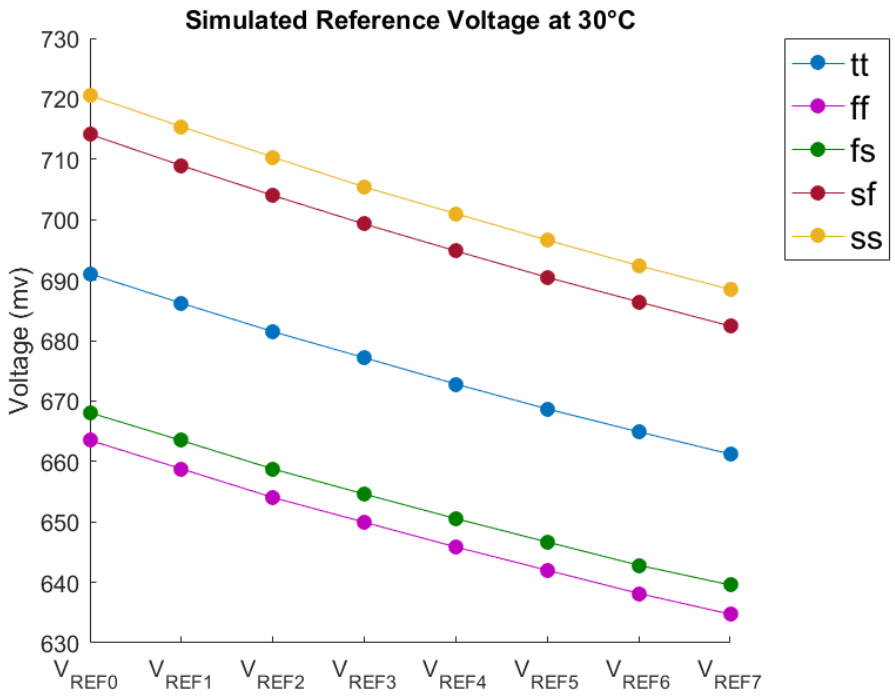

Figure 4.4: Simulated Reference Voltage Output over Process Corner at 30°C

This shows that there is a wide range of output voltages between corners, and each corner has an output voltage decreasing mostly linearly for the different reference voltage circuits. For example, fast-fast and slow-slow show about a 50mV difference. The typical

corner is right in between two groups, one consisting of the nMOS slow corners and then the nMOS fast corners. There is only a difference of about 8mV between the fs-ff corners and then the sf-ss corners, indicating that the nMOS corner is the dominating effect on the reference voltage output and the pMOS corner effect is minimal, which makes sense considering the equation for the output voltage is a function of both an nMOS threshold voltage and an nMOS k parameter.

The total temperature variation is plotted in figure 4.5.

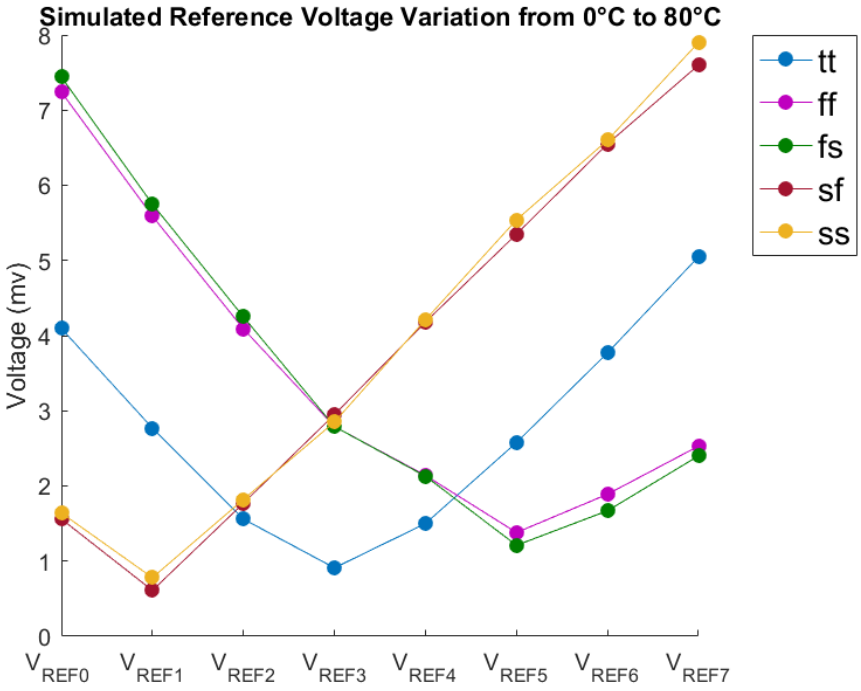

Figure 4.5: Simulated Voltage Variation Across Temperature and Process Corner

This shows that the maximum variation over process is about 8mV, and each process will have a reference voltage with variation less than 2mV, so that voltage reference circuit with the least variation could be used as the reference voltage for the entire

application on chip. This plot shows that the corners, ff and fs are nearly identical from a variation across temperature perspective. The same is true for ss and sf. This means essentially, there are three corners to consider: typical, nMOS slow, and nMOS fast. The slow corners have their minimum variation at V_{REF1}, and the fast ones have their minimum at V_{REF6}. The typical corner has its minimum at V_{REF3}, which is also where the slow and fast corners intersect. The big difference in V_{REF} values from the previous graphs show that in production, the circuit would also need trimming if an exact bias voltage is needed. In particular, the threshold voltage is very process-dependent, which leads to a wide-range of reference voltage values. The average value of the measured reference voltage across temperature is 675.88mV, and the average variation across temperature over the five corners is 3.49mV.

4.5 LAYOUT

The design was laid out in the 180nm TSMC process node. Figure 4.6 shows the layout for a single voltage reference circuit, and figure 4.7 shows all 8 voltage reference circuits on the final chip. The area for a single circuit is 64µm x 64µm, or 4.096nm^2. The area for all 8 circuits together is 62.4nm^2. The majority of the area is used for the BJT devices because there are eleven total BJTs, but an extra non-connected BJT cell is added to make the layout symmetrical, for a total of twelve BJTs in each circuit. The capacitor is implemented as a MIM (metal-insulator-metal) capacitor.

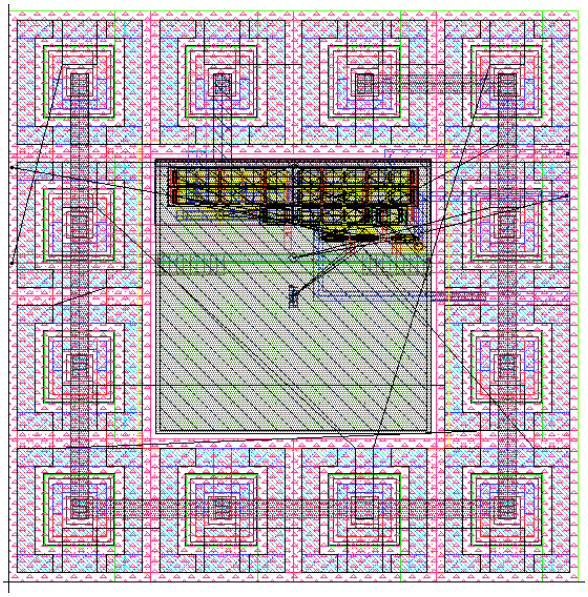

Figure 4.6: Layout of Single Voltage Reference Circuit

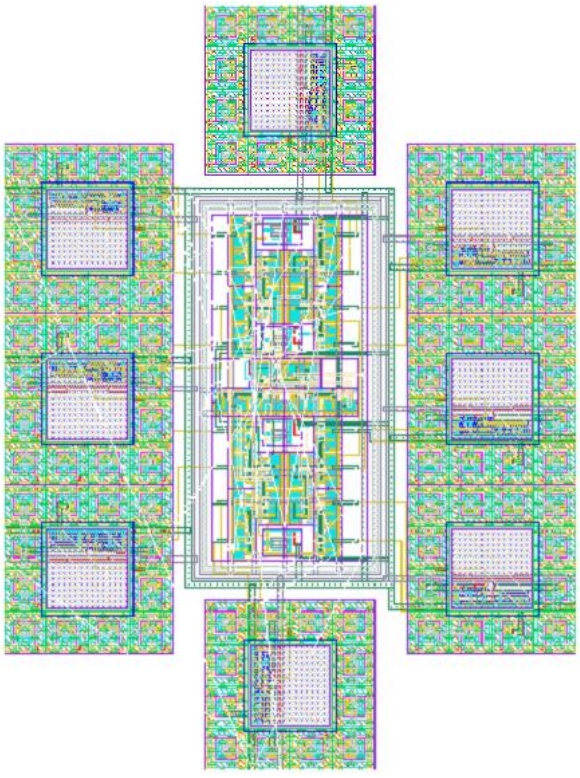

Figure 4.7: Layout of All Voltage Reference Circuits

In the layout for the all eight of the circuits, the voltage references are on the edges and the middle block is the digital block used for muxing the chosen voltage reference to the test pin.

Chapter 5: Measurement of Simple Voltage Reference Circuit

The primary measurement necessary to take for a voltage reference circuit is the voltage measurement over a range of temperatures. This allows the actual voltage reference drift over temperature to be measured. This measurement requires only a DC multimeter, DC power supply, and temperature chamber. This chapter details the test setup, test equipment, measurements, and then extra Monte Carlo simulation data to help reconcile the measured results with the simulation results.

5.1 Test Setup

This section details the setup used for testing the chips. The chip is placed into socket on the test board, which is then placed into the temperature chamber. Wires from the power supply and DC multimeter are connected to the test board. In order to access the voltage reference circuits on the chip, the correct digital code must be passed through the scan chain on the front end of the chip. This will select the appropriate circuit from the test mux. The digital I/O levels on the chip are at 1.8V. In order to send in a bit stream and clock to select the circuit, a Raspberry Pi was used. This uses Python scripts to control the general purpose input/outputs (GPIO). The only issue with this is that the Raspberry Pi has GPIO levels of 3.3V. Therefore, a level shifter is used to shift the 3.3V level down to 1.8V. To implement the level shifter, a Xilinx CPLD is used. It is easy to program with simple Verilog Code. Once the correct digital stream has been passed into the scan chain, the temperature is changed from 0°C to 80°C in steps of 10°C. For IC

temperature testing, it is important for the IC to soak at the temperature set on the temperature chamber for a period of time to ensure that the whole die is at the desired temperature. Essentially, it takes time for the whole IC to reach the temperature set by the chamber. For this test, a soaking time of 6 minutes was used at each temperature. A longer soaking time of 10 minutes was used for the initial temperature drop of room temperature to 0°C because it is a larger temperature step than the usual 10°C step used for the rest of the testing.

5.2 TEST EQUIPMENT

This section details the equipment and part numbers used in the testing. The multimeter is a 6-and-a-half digit precision multimeter, and the power supply has three output supply voltages.

Item	Manufacturer	Part Number
DC Power Supply	Agilent	34401A
DC Multimeter	Agilent	E3631A
Temperature Chamber	Associated Environmental Systems	SD-502
TQFP-44 IC Socket	Yamaichi Electronics	IC51-0444-467
CPLD	Xilinx	XC2C64A
Raspberry Pi	Raspberry Pi	RPi3 Model B

Table 5.1: Test Equipment used for Measurements

5.3 Measurement Results

This section details the actual measurement results. The measurements were recorded from 0°C to 80°C for five chips, and eight voltage reference circuits per chip. A plot is made similar to the one in the simulations section, in which the magnitude of the reference voltage at 30°C is plotted for each reference voltage circuit on each chip.

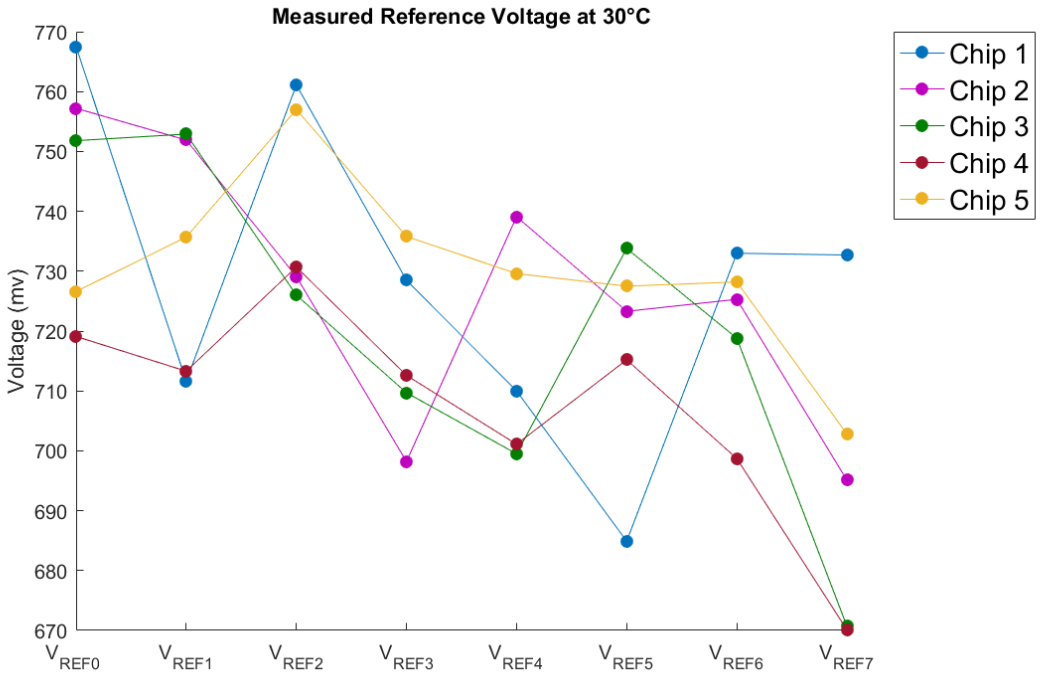

Figure 5.1: Measured Output Voltage at 30°C

The most obvious difference between the measured data and the simulated data is that predominately, the output voltages are much higher than the simulation. The average value of the measured output voltage was 723.3mV, compared to 675.88mV for simulation. Also, the reference voltage does not follow a linear slope as the simulation

results did. There are many instances of voltage increasing from one reference voltage circuit to the next, even though in simulation, it is monotonically decreasing for each reference voltage circuit step. It is seen though, that overall, there is a decrease in voltage going from V_{REF0} to V_{REF7}. Because of the high magnitude of the reference voltage, the chips measured are likely slow-slow or slow-fast parts.

The next plot is a plot of the output voltage variation across temperature for the different circuits. The variation was found by finding the maximum and minimum measurements for a specific voltage reference circuit on a specific chip and then computing their difference. The format of this plot is just like the one done in the simulation section.

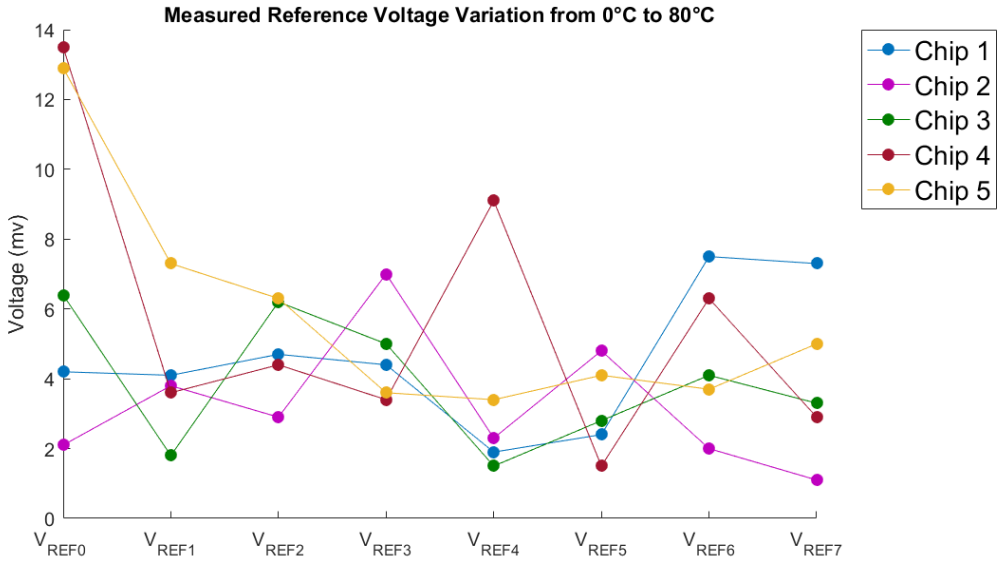

Figure 5.2: Measured Output Reference Voltage Variation across Chip and Temperature

The plot of the variation shows that over five chips, each reference voltage circuit has an instance of being less than 4mV variation over an 80°C range. Furthermore, each

chip has at least one voltage reference circuit that has less than 4mV variation. This means that although there are a few circuits with a relatively high variation of about 13.6mV, that circuit does not have to be used and the best (lowest variation) reference can be chosen to be the reference for that part of the chip. The average temperature variation is 4.62mV, compared to 3.49mV for the simulation across process corners.

In the simulation, the best-performing reference voltage circuit for the slow nMOS corners was V_{REF1}, but in this circuit, the best performing circuit for three of the five chips is V_{REF4}, and then V_{REF5} and V_{REF7} have the least variation for the other chips. The next plot is a plot of the best-case voltage variation for each chip across temperature.

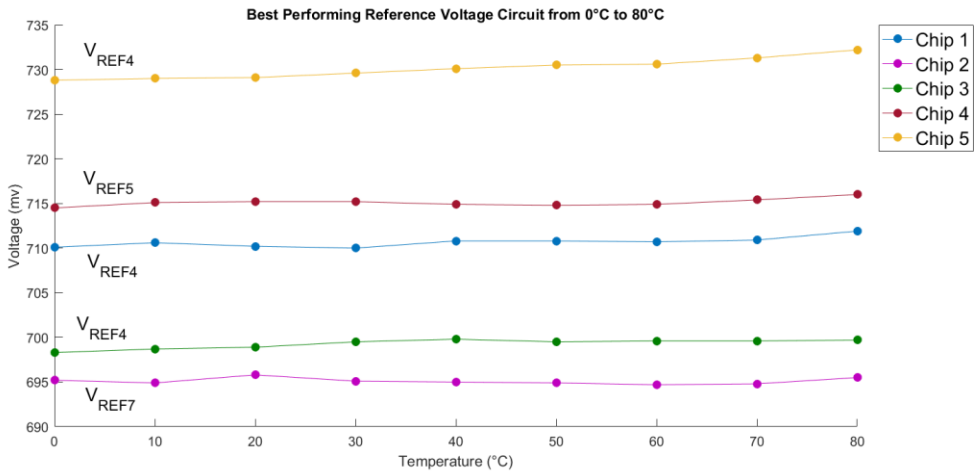

Figure 5.3: Best Performing Reference Voltage Circuit from 0°C to 80°C

Figure 5.3 shows that the voltage reference circuit with the best performance per chip can have a wide difference in voltage. The difference is about 35mV across the five chips. For a 1.8V power supply, this is a difference of about 2% of the full-scale voltage, which is not very significant. In a data converter, if the reference voltage used is coming

from this circuit, that means that each code will just correspond to a shifted voltage, but in communication systems, that will not make a major impact. In a precision measurement system, this performance would be very poor.

5.4 Monte Carlo Simulation Results

The purpose of this section is to show, from simulation, the reason for why there is such a wide-range of performance across the five chips measured. Prior to tape-out, the primary simulation was done over the process corners. Device mismatch, which is simulated by using Monto Carlo simulation, was thought to be negligible because the devices use a long channel of 2µm. The widths of the nMOS devices are minimum in the case of the n-channel transistors which act as inputs for the two base-emitter voltages. Even though these widths are small, the total device area is large, so mismatch was thought to have minimal effect on the circuit performance. From the process corner simulations, the simulated results from chapter 4 show the reference voltage values can be predictable based on the process corner. However, the results in this chapter, such as in figure 5.1, show that there is a very wide variance in the reference voltage measured across the same circuit design across chips, so it is difficult to choose the reference circuit to use based on just the magnitude of the reference voltage measured at room temperature.

Because of these issues, possible reasons for this measurement-simulation mismatch were considered. This led to the Monte Carlo simulation which applies device mismatch to a group of selected devices and produces a histogram of the results. In these simulation

results presented, one hundred simulation runs were specified, and instead of just taking the histogram, the reference voltage variation over temperature (0°C to 80°C) is plotted against the reference voltage magnitude at 30°C. By plotting in this way, it enables the voltage variation over temperature to be seen for a particular reference voltage magnitude. The measured values were, on average, high, so they are most comparable to the simulation results for the slow-slow corner, so the simulations here are done for the slow-slow corner.

The first devices considered for the mismatch were the input nMOS devices which sense the base-emitter voltages. This plot is shown in figure 5.4 for the first reference voltage circuit (V_{REF0}) with a diode-connected device width of 360nm. This shows that the mismatch between the input devices is enough to lead to a very wide range of reference voltage values and the performance over temperature. There is a difference of about 80mV in the 30°C value over the one hundred simulation runs, but the majority lie between 700mV and 765mV, which is consistent with the measured results for V_{REF0}.

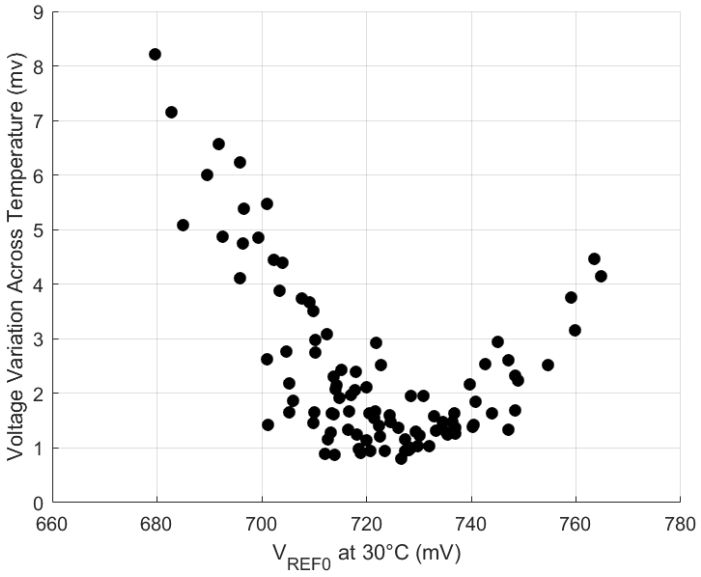

Figure 5.4: Monte Carlo Results for V_{REF0} with Input nMOS Device Mismatch Only

Then, the nMOS diode-connected output device was considered by itself. Its plot is shown in figure 5.5, and there is about a 20mV range on the reference voltage value at 30°C, and the variation over temperature is always less than 1.7mV. Therefore, the measurement difference in the actual silicon versus the prior simulation is not caused by device mismatch of this output device.

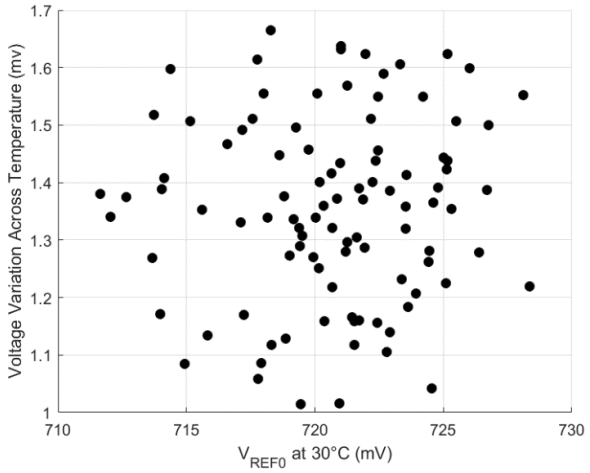

Figure 5.5: Monte Carlo Results for V_{REF0} with Output nMOS Device Mismatch Only

Next, the PMOS devices only were considered. There is significant variation in the results due to these devices' mismatch. There is about a 50mV difference in the reference voltage at 30°C, and more variation in the performance across temperature.

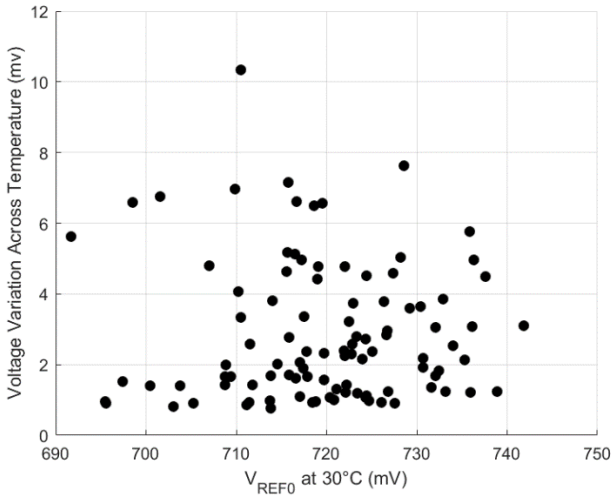

Figure 5.6: Monte Carlo Results for V_{REF0} with pMOS Device Mismatch Only

As shown, there is significant variation due to the mismatch as shown in the Monte Carlo simulation. The biggest factor is the n-channel input devices, and the pMOS

mismatch also contributes significantly to chip-to-chip variation. From the process corner simulations, the slow-slow reference voltage values are the highest compared to all other corners, and the reference voltage circuit which has the lowest reference voltage is V_{REF7} which has an output device width of 500nm. Because this should ideally have the lowest voltage, but the measurement results of chip 1 show a reference voltage value at room temperature of about 735mV, so Monte Carlo Simulation was run considering all nMOS and pMOS devices to have a mismatch to see how much variation is predicted by the simulation. The plot is shown in figure 5.7. It does not show a value as high as 735mV, but it is close to 730mV, so it does confirm that even though V_{REF7} should have a lower value, because of the mismatch, there is significant variation, with values ranging from about 653mV to 728mV, and can make the V_{REF7} output value much higher than initially predicted by process corner simulation alone.

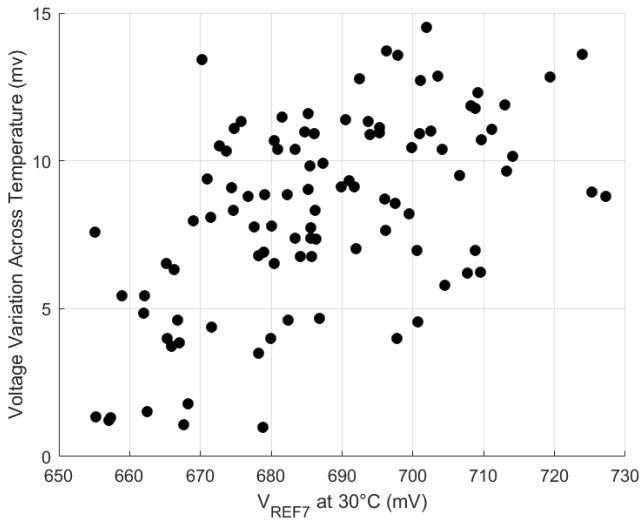

Figure 5.7: Monte Carlo Results for V_{REF7} with All MOS Device Mismatch

The Monte Carlo simulation results show that, compared to the process corner simulations alone, there are major variations in the performance of the reference voltage circuit and the nominal reference voltage value. The reference voltage value could be adjusted to the desired value, but there is still the issue of knowing which reference voltage circuit to use on a given chip which is fabricated in a particular process corner. The circuits, as they exist now, are not suitable for rapid trimming because to know which reference circuit to use, it will require a temperature sweep, and even when a group of chips are at the same process corner, some chips may have a different optimum voltage reference circuit due to mismatch, as shown by the Monte Carlo simulation. One good aspect of this voltage reference circuit is that by testing at the lowest temperature and then the highest temperature should give an accurate indication of the best performance voltage reference circuit, so at least the number of temperature steps during the sweep can be reduced to the minimum and maximum temperatures.

Chapter 6: MOSFET-Only Simple Voltage Reference

In this chapter, a variation on the designed simple voltage reference is presented in which the BJTs are replaced with diode-connected nMOS devices. This innovation leads to a reduction in circuit area. The proposed circuit is designed in the TSMC40nm process with a nominal supply voltage of 1V, which is the operating condition for which a sub-1V reference voltage would be desired. The results presented here also demonstrate better performance across process corner and device mismatch. With these benefits, combined with the reduction in area, this new circuit has multiple advantages over the previous implementation with BJTs. The results presented in this chapter come from simulation using Cadence IC 6.1.6 and the models are based on BSIM4. The sections of this chapter include an analysis of temperature variation for diode-connected MOS devices, an overview of the proposed circuit, and an analysis of simulation results including Monte Carlo simulations.

6.1 Temperature Variation for Diode-Connected MOS Devices

As shown in chapter 4, the threshold voltage for a MOS device is CTAT, and as shown in chapter 2, the difference between two CTAT voltages is PTAT. For a constant current in a device, the overdrive voltage, which is the difference of the gate-source voltage and the threshold voltage, will be CTAT. The sample circuit used to simulate this is shown in figure 6.1.

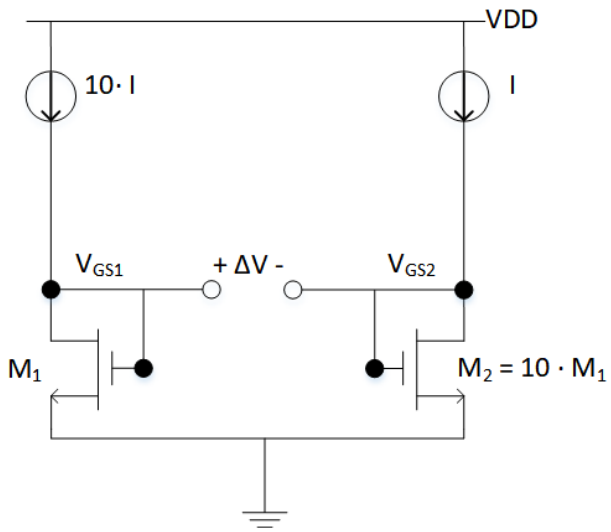

Figure 6.1: Sample Circuit for PTAT Generation Using MOSFETs

For this circuit, the current in M_1 is ten times that of M_2.

$$I_{D1} = 10 * I_{D2} \tag{6.1}$$

Substituting the long-channel model equation for the drain currents, and writing the current scaling factor as *m*,

$$k * {V_{OV1}}^2 = m * k * {V_{OV2}}^2 \tag{6.2}$$

Here, V_{OV} is the overdrive voltage. Then, cancelling the transconductance parameter and taking the square root,

$$V_{OV1} = \sqrt{m} * V_{OV2} \tag{6.3}$$

Lastly, re-writing equation (6.3) as the difference in over-drive voltages,

$$V_{OV1} - V_{OV2} = V_{OV2} * (\sqrt{m} - 1) \tag{6.4}$$

The circuit in figure 6.1 was simulated in cadence and the difference in gate-source voltages is plotted in figure 6.2.

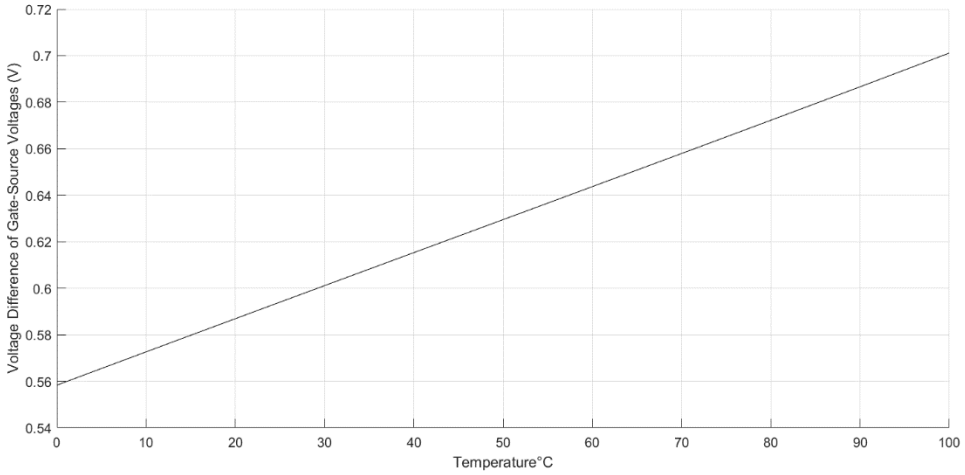

Figure 6.2: PTAT Voltage Generated by Difference of Gate-Source Voltages

Figure 6.2 shows that the voltage difference between the gate-source voltage of two diode-connected nMOS devices is indeed PTAT. Therefore, the BJTs in the circuit proposed in chapter 3 can be replaced with MOSFETs, which will provide a reduction in area for the circuit.

6.2 Simple Voltage Reference Using MOSFETs Only

The architectural change to the circuit introduced in chapter 4 is simply replacing the BJTs with diode-connected nMOS devices. This is shown in figure 6.3.

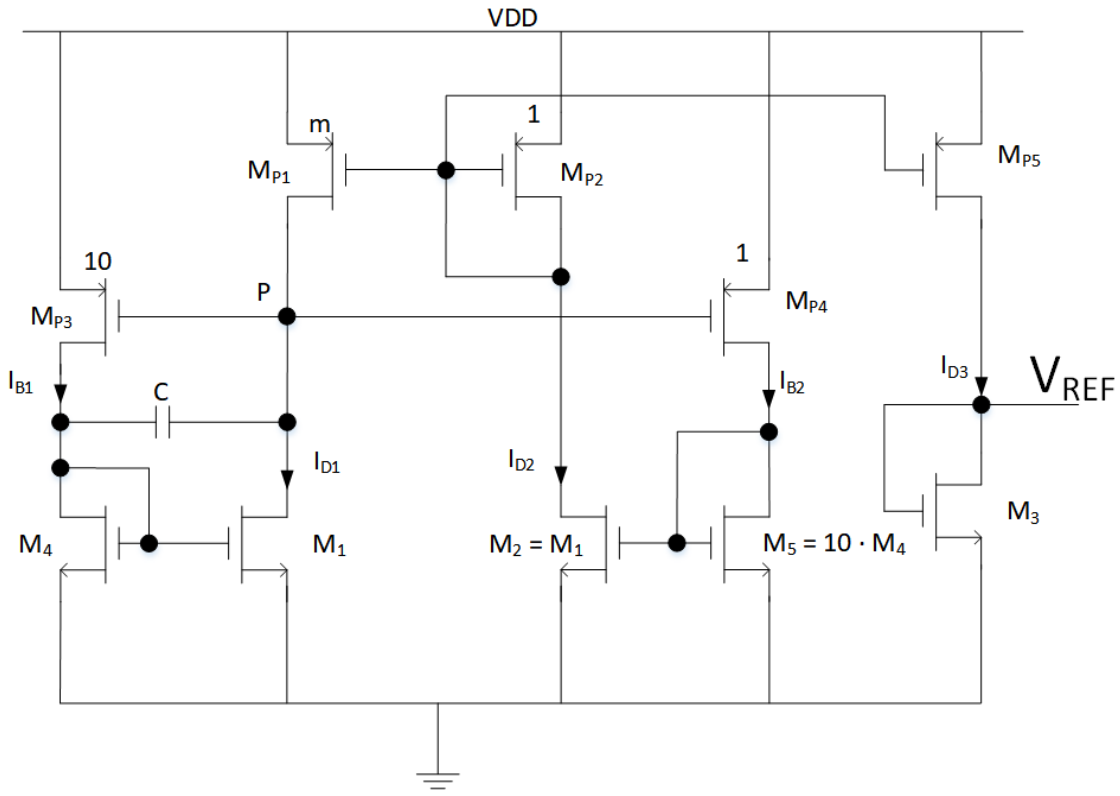

Figure 6.3: Simple Voltage Reference Circuit Using MOSFETs Only

The circuit has the same fundamental equations for operation as in chapter 4, except for the two inputs to the differential pair now come from diode-connected devices.

An important consideration used in this circuit is that for M_1 and M_2 to turn on, these devices must be low threshold voltage devices, whereas M_4 and M_5 will be regular threshold voltage devices. Furthermore, the output device, M_3, can be replaced with different threshold voltage devices to achieve different levels of reference voltage.

6.3 Simulation of MOSFET-Only Reference Voltage Circuit

The circuit was designed in the TSMC40nm process and uses a 1V supply voltage. This makes it more realistic for an actual implementation of a sub-1V voltage reference because the primary reason for using this is due to the shrinking supply voltage which makes a standard bandgap voltage of 1.2V unrealizable. The results presented here show a higher nominal value of the reference voltage and a more constant voltage reference over temperature considering mismatch. The circuit which was simulated uses a low threshold voltage device at the output to achieve a reference voltage less than 800mV because of the higher nominal reference voltage values generated in this circuit.

The simulation was done over process corners and temperatures, and mismatch using Monte Carlo simulations to see the performance with device mismatch considered. In the prior circuit which uses BJTs, mismatch was a major issue, which not only affected the nominal value of the reference voltage, but also greatly affected the variation across temperature. This made it difficult to obtain a stable reference voltage on the chip which was taped out without sweeping all reference voltage circuits across temperature.

The following graphs show the reference voltage across temperature, process corner, and device mismatch, where the five same corners as before (tt, ff, fs, sf, ss) are considered. Unlike the circuit using BJTs, this is for the same design without changing any transistor sizes. Before the graphs are shown, a table is presented showing the statistics from the Monte Carlo simulations, which use one hundred samples.

Process Corner	Min Variation	Max Variation	Mean	Std. Deviation
Typical	1.157mV	4.115mV	1.735mV	.521mV
Fast-Fast	1.220mV	4.237mV	2.240mV	.674mV
Fast-Slow	1.208mV	3.284mV	1.808mV	.487mV
Slow-Fast	1.087mV	5.452mV	2.125mV	.833mV
Slow-Slow	1.002mV	6.107mV	2.63mV	.986mV

Table 6.1: Statistics for Monte Carlo Simulation for MOSFET-Only Reference Circuit

Figure 6.4 shows the variation over temperature (0°C to 100°C) versus the reference voltage value at room temperature (27°C) for the typical process corner.

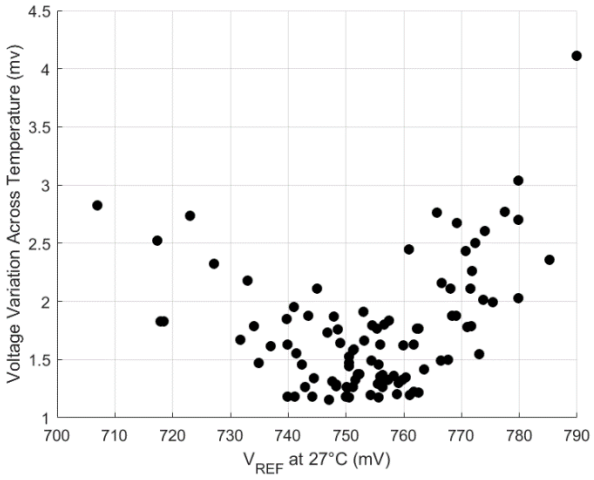

Figure 6.4: Monte Carlo Results for Typical Corner

This plot shows a maximum voltage variation of about 4.1mV, with the vast majority having a variation of less than 3mV. The lowest variation values come for a nominal reference voltage value between 740mV and 760mV. By trimming the transconductance parameter for the output pMOS device, the reference voltage value can be trimmed to be in this range, which would give a stable reference voltage across temperature. The benefit

of this circuit is that even though the reference voltage value at room temperature changes over a wide range, the voltage variation over temperature is still low.

The next figure shows the Monte Carlo results for the fast-fast process corner. Similarly to the typical corner, the lowest variation values come for a nominal reference voltage value between 740mV and 760mV, and the relationship between the variation across temperature to the nominal value is mostly linear. Because the nMOS mismatch has the dominant effect on the mismatch, the fast-slow corner is very similar, so it is not plotted in this thesis.

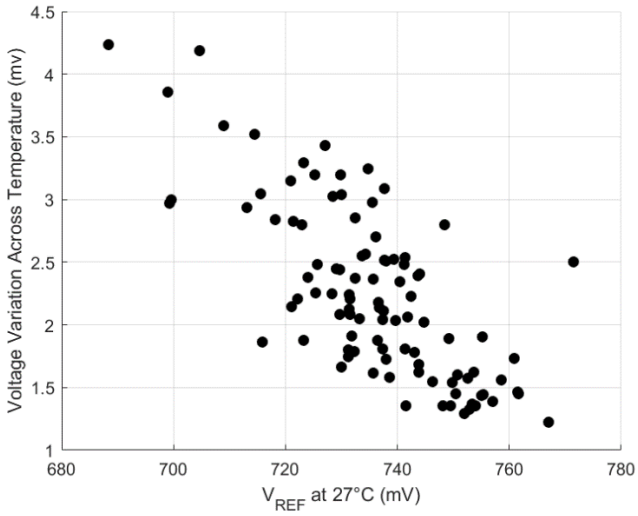

Figure 6.5: Monte Carlo Results for Fast-Fast Corner

Next, the slow-slow corner Monte Carlo simulation results are shown. This has more variation across temperature, and has an inverse slope compared to the fast-fast corner. This also has low voltage variation between 740mV and 760mV.

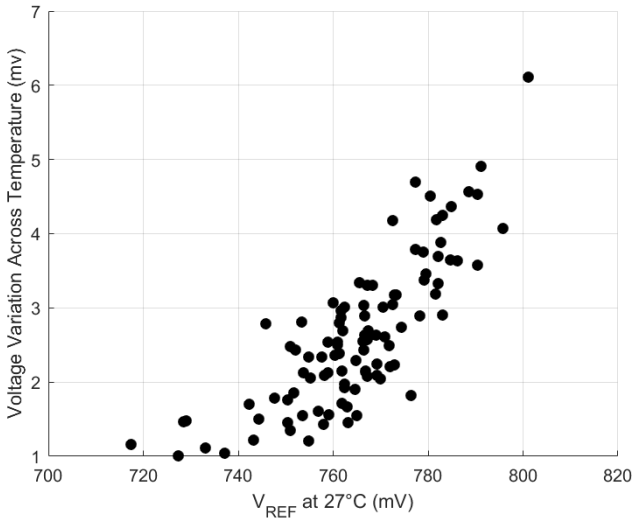

Figure 6.6: Monte Carlo Results for Slow-Slow Corner

Together, the Monte Carlo simulation results show that even without changing any device sizing, the voltage reference variation across temperature is decreased significantly from the voltage reference circuit using BJTs. Furthermore, the results show that between 740mV and 760mV, the voltage variation across temperature is almost always under 3mV. This indicates that if the transconductance parameters of the output devices, either the pMOS current mirror or nMOS output diode-connected device, can be trimmed to this range of values, then the circuit should produce a stable voltage reference across temperature.

The next step for this circuit would be to do a full layout, run the parasitic extraction, re-simulate to make sure parasitic effects do not degrade the circuit performance, and then tape it out. It is expected that because there is less variation that results from mismatch, then the results should be closer to simulation, as opposed to the reference voltage circuit with the BJTs which was previously taped-out. This circuit also take less

area than the first circuit because the diode-connected nMOS devices replacing the BJTs have an area of 240µm by 1µm, as opposed to 2µm by 2µm for the BJTs. Also, by using TSMC40nm, the other MOSFETs will also have less area.

Chapter 7: Conclusion

This work presented a simple design for a sub-1V voltage reference circuit, as well as some important and fundamental implementations of bandgap voltage reference circuits from the past. The simulation results and actual measurements from the chips containing the circuit were also shown. The key motivation for a sub-1V voltage reference is that as supply voltages are decreasing, it is no longer possible to achieve the full 1.2V from a standard bandgap circuit. Also, many bandgap circuits require the use of a high-gain, low voltage offset op amp which, in and of itself, can be a complicated design problem. The circuit demonstrated in this paper is simple in comparison to many other implementations and requires no voltage divider circuit, so no resistors are needed, nor does it require a third BJT. The proposed circuit also does not need a start-up circuit which is another advantage of not having an op amp in the circuit since the op amp's start-up output condition is unknown which makes prior implementations necessitate a start-up circuit. The circuit is novel in its generation of the PTAT component by using a differential pair with the base-emitter voltages as inputs to the differential pair, and then a feedback mechanism for reaching a steady state where the drain currents of the differential pair are PTAT. The CTAT voltage also does not come from a BJT, but instead comes from the threshold voltage of a MOS device. The final output reference voltage comes from passing the PTAT current through the diode-connected device. The voltage reference circuit itself is also small, so multiple copies can be placed on a chip, with each circuit optimized for a particular process corner.. The achieved performance of this circuit shown in chapter 5 shows that each of the five measured chips have at least

one reference circuit that can produce an output voltage with a voltage variation less than 4mV across an 80°C temperature sweep, thus showing it can be useful in applications such as a 6-bit data converter system. Further trimming will be required to achieve a specific voltage reference value, and low-threshold or high-threshold voltage devices can be added into the circuit to decrease/increase the reference voltage value.

The "trim" method described in this thesis was to make multiple copies of the circuit with a wide enough range of device sizing such that depending on the process corner, a specific reference voltage circuit would be chosen. The measurement results show that there is a wide variation over the output voltage value, and there is also variation across chips between which reference voltage circuit has the lowest voltage variation over temperature. There is significant process variation that is affecting the manufactured IC that is coming from the mismatch between devices. This is shown by the Monte Carlo simulation results which indicate significant chip-to-chip variation.

An alternative implementation of the circuit replacing the BJTs with diode-connected MOSFETs was presented using simulation results which indicate less chip-to-chip variation. This design is also implemented in the TSMC40nm which would be an ideal process for a sub-1V reference voltage because the typical 1.2V bandgap voltage is not attainable with a 1V supply. This circuit also has a significant area reduction over the circuit with BJTs. Further work must be done on this circuit to tape it out and measure actual silicon performance, although the improved performance from Monte Carlo simulation shows that this circuit could have consistently low voltage variation even across process and mismatch considerations.

Appendix

In CMOS-only processes, it is not possible to do a full BJT layout as would be available in a Bi-CMOS process. This issue is overcome by using substrate pnp, or vertical, devices. The device cross section is shown in the figure below [7].

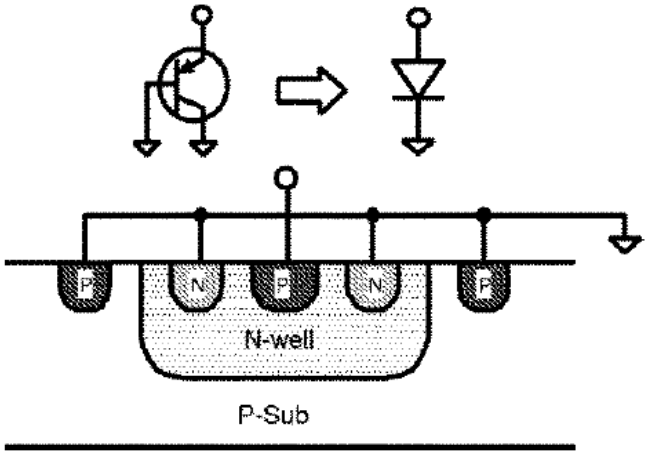

Figure A.1: Cross-Section of Substrate pnp Devices in CMOS-only Process

The pnp structure is formed by the P-TAP in the N-Well, the N-Well, and the P-Substrate. By grounding both the N-Well and the P-Substrate, the equivalent circuit of the substrate pnp device is a pn junction, which is just a diode. Thus, in this thesis, whenever a diode is drawn, it is implying this pnp substrate implementation of a diode. It is noted that the equation for the forward voltage drop across a diode is equivalent to that of a bipolar base-emitter voltage. This is written below, where the V_F is the forward voltage, and I_D is the diode current, and the other terms have their usual meanings as in earlier sections.

$$V_F = \frac{kT}{q} \ln\left(\frac{I_D}{I_S}\right) \tag{A.1}$$

References

[1] R. J. Widlar, "New developments in IC voltage regulators," *IEEE Journal of Solid-State Circuits*, vol. 6, pp. 2–7, Feb. 1971.

[2] D. Feucht, “Signal-Processing Circuits”, in *Handbook of Analog Circuit Design*. San Diego, CA: Academic Press, 1990, ch. 11, sec. 11.1, pp 525.

[3] B. Razavi, “The bandgap reference,” *IEEE Solid-State Circuits Mag*., vol. 8, pp. 9–12, Summer 2016.

[4] A. P. Brokaw, “A simple three-terminal IC bandgap reference,” *IEEE J. Solid-State Circuits*, vol. 9, no. 6, pp. 388–393, Dec. 1974.

[5] R. Gregorian, G. Wegner, and W. E. Nicholson, “An integrated single-chip PCM voice codec with filters,” *IEEE J. Solid-State Circuits*, vol. 16, no. 4, pp. 322–333, Aug. 1981.

[6] H. Neuteboom, B. M. Kup, and M. Janssens, “A DSP-based hearing instrument IC,” *IEEE J. Solid-State Circuits*, vol. 32, pp. 1790–1806, Nov. 1997.

[7] H. Banba, H. Shiga, A. Umezawa, and T. Miyaba, “A CMOS bandgap reference circuit with sub-1-V operation,” *IEEE J. Solid-State Circuits*, vol. 34, no. 5, pp. 670–674, May 1999.

[8] G. Groesenken et. al. "Temperature Dependence of Threshold Voltage in Thin-Film SO1 MOSFET's", *IEEE Electron Device Letters*, Vol. 11, pp. 329-332, August 1990.

[9] F. P. Heiman and H. S. Miller, “Temperature dependence of n-type MOS transistors,” *IEEE Trans. Electron Devices*, vol. ED-12, no. 3, pp. 142–147, Mar. 1965.

Vita

Peter Lisle Schaeffer was born in San Antonio, TX, and grew up in Seguin, TX and New Braunfels, TX. He graduated from Lifegate Christian School in Seguin, TX, in May 2012. Starting in the fall of 2012, he began his undergraduate degree at the University of Texas at Austin in the Department of Electrical and Computer Engineering. Along the way, he also took enough courses in the Asian Studies Department to be awarded a minor in Asian Studies. He was accepted into the Integrated BSEE/MSE program in the summer of 2015, and began full-time graduate school in the fall of 2016. He served as a Graduate Teaching Assistant and received the "Best TA" award. Upon graduation with his Masters degree, he plans to join Arm in Deerfield Beach, FL, where he will work on the design of analog/RF integrated circuits for the Internet of Things group.

Address: p.lisle.schaeffer@gmail.com

This thesis was typed by the author.